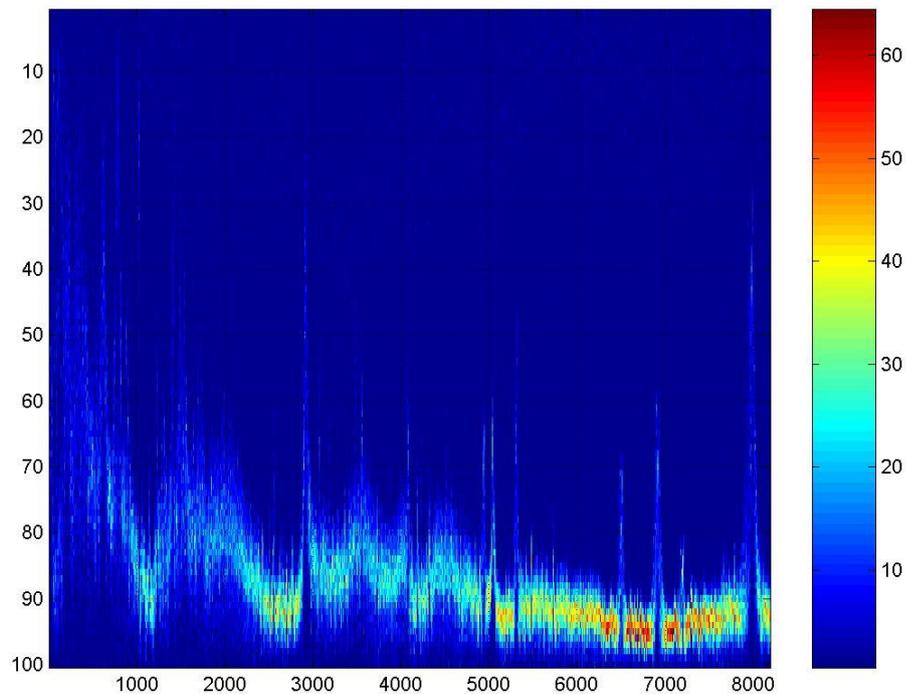


Sergio Frasca

Laboratorio di Segnali e Sistemi II

Dipartimento di Fisica
Sapienza Università di Roma



Versione 25 febbraio 2019

Versione aggiornata in <http://grwavsf.roma1.infn.it/Issii/Issii.pdf>

Contents

1 - Introduzione al corso.....	7
2 - Segnali discreti	8
Tipi di segnali	9
Segnali notevoli.....	10
3 - Segnali transitori.....	13
Parametri che caratterizzano un segnale transitorio	13
Parametri che caratterizzano due impulsi	16
Trasformata di Fourier di un segnale ad energia finita	17
Autocorrelazione e spettro di energia	19
4 - Trasformate	20
Trasformata z	20
Analogia con la trasformata di Laplace	20
Proprietà della trasformata z	23
Alcune trasformate z	25
Trasformata di Fourier	26
La trasformata di Fourier per dati discreti (DTFT)	35
La DFT (e la FFT)	37
Trasformata di Hilbert e segnale analitico	38
Altre trasformate.....	41
5 - Sistemi di acquisizione.....	42
Schema di un sistema di acquisizione.....	42
Teorema del campionamento e aliasing	44
Generalizzazione.....	47
Il caso dei segnali complessi	48
Il rumore di quantizzazione.....	49
Signal conditioning.....	51
6 - Sistemi discreti e modelli	53
Generalità sui sistemi	53
Classificazione dei sistemi.....	56
Sistemi discreti	59

Casi particolari.....	61
Semplice applicazione di sistemi MA e AR.....	62
Equazioni alle differenze.....	65
Funzione di trasferimento discreta.....	67
Differenze e derivate.....	69
Risposta di un sistema discreto.....	72
Risposta impulsiva	72
Risposta forzata	73
Evoluzione libera.....	73
Stabilità	75
Sistemi semplici	76
Sistema di ordine 0	76
Sistema MA del primo ordine	76
Due sistemi MA del primo ordine in cascata; MA del secondo ordine	84
Sistema AR del primo ordine (reale).....	87
Sistema AR del primo ordine (complesso)	96
Sistema AR del secondo ordine	100
Semplici sistemi ARMA	104
Sistemi non-lineari - i sistemi di Volterra	109
7 - Segnali aleatori.....	110
Introduzione ai processi stocastici.....	110
Definizione di processo stocastico	110
Funzioni del secondo ordine	111
Il caso di due processi	113
Stazionarietà, ergodicità	114
Esempi	116
Processi normali.....	118
Il rumore bianco.....	119
Trasformazioni di processi stocastici.....	120
Sistema statico (senza memoria).....	120
Sistema lineare (tempo invariante).....	121
Un caso particolare: il derivatore.....	123

Processi stocastici normali	125
Proprietà fondamentali	125
Processi stocastici discreti	127
Rumore bianco discreto.....	128
Processi normali e sistemi discreti	131
Résumé dei risultati	132
Processi AR, MA e ARMA	135
Processo AR del primo ordine (reale)	135
Processo AR del primo ordine (complesso).....	137
Processo AR del secondo ordine.....	140
Processo di Poisson	142
8 - Analisi statistica dei segnali.....	144
Caratteristiche statiche	144
Autocorrelazione.....	146
Spettro di potenza	149
Stimatori spettrali non parametrici	150
Cross-correlazione e cross-spettro	165
Coerenza.....	167
9 - Filtraggio	168
Segnali e rumori, rapporto segnale/rumore.....	168
Il filtro adattato	171
Caso del rumore bianco	172
Altre dimostrazioni	175
Caso generale	177
Teoria della rivelazione (Detection Theory).....	180
Filtro di Wiener.....	185
Realizzazione di filtri	187
Filtri FIR.....	187
Filtri IIR	187
Filtri a sfasamento nullo	187
Filtri in frequenza	189
Filtri adattivi.....	190

10 - Cenni all'analisi delle immagini.....	191
Immagini ed elaborazione delle immagini.....	191
Elaborazione lineare delle immagini	195
11 - Cenni alla Teoria dell'Informazione	204
Schema di trasmissione di un segnale.....	204
Entropia e informazione	205
Compressione	206
Codici a correzione	207
12 - Introduzione all'apprendimento automatico (Machine Learning)	208
Appendice - Trasformate z.....	209

1 - Introduzione al corso

Con lo sviluppo e la diffusione dei computer il modo di fare misure in fisica e nella tecnica è sostanzialmente cambiato: molto spesso ora il risultato di una misura si ottiene con l'acquisizione di dati campionati e la loro elaborazione. In questo modo si ottengono notevoli vantaggi, tra cui, in genere, anche un più basso costo.

Per elaborare i dati campionati si sono sviluppate particolari tecniche matematiche e statistiche.

In questo corso ci occuperemo di acquisizione ed elaborazione di segnali digitali per le misure in fisica.

Le grandezze fisiche in genere variano in modo continuo nel tempo.

Per misurarle usiamo dei trasduttori che le trasformano in grandezze elettriche (molto spesso tensioni elettriche).

Queste tensioni elettriche sono quindi "campionate" e convertite in formato digitale, ottenendo una successione di numeri a cui ci riferiremo come segnali digitali. Questi sono memorizzati in un computer o su opportuni supporti (acquisizione dei dati), dove possiamo elaborarli convenientemente.

Nel seguito vedremo come ottenere "buoni" segnali digitali, come rappresentarne le proprietà, quali sono le tecniche base per estrarre informazioni di nostro interesse.

Accenneremo anche al caso in cui i dati sono immagini digitali e a metodi di elaborazione dei dati basati sulle reti di neuroni artificiali.

Il corso e le esercitazioni di laboratorio si avvalgono dell'uso di Matlab, un potente ambiente software e linguaggio per l'elaborazione numerica.

Molti esempi scritti in Matlab complementano le lezioni teoriche.

2 - Segnali discreti

Live Script [segnali.mlx](#), [energyspec.mlx](#), [periodici.mlx](#)

Un segnale è una funzione del tempo¹ che rappresenta una grandezza fisica.

Un segnale è detto a tempo continuo, o semplicemente continuo, se è definito per tutti i valori di un intervallo della variabile reale tempo² (t); per esempio una tensione elettrica

$$(2.1) \quad x(t) \quad \text{per } t_{\text{iniziale}} \leq t \leq t_{\text{finale}}$$

È detto discreto se è definito per un insieme discreto dei valori della variabile t; in tal caso in genere viene rappresentato dalla successione di questi valori

$$(2.2) \quad x_1, x_2, \dots, x_N$$

$\{x_1, x_2, \dots, x_N\}$ è detta successione o serie temporale.

Talora gli indici non partono da 1, ma possono prendere anche valori negativi (e ovviamente 0). Spesso un segnale discreto viene ottenuto “campionando” un segnale continuo, cioè estraendo i valori del segnale continuo da un insieme discreto di valori del tempo t_1, t_2, \dots, t_N ; considerando il segnale continuo $x(t)$, otteniamo i valori

$$(2.3) \quad x(t_1), x(t_2), \dots, x(t_N)$$

Tali valori sono detti campioni (samples) e possono essere indicati come nella (2.2). Tale procedura viene detta campionamento (sampling); particolarmente usato è il campionamento uniforme, in cui è costante la differenza $\Delta t = t_{i+1} - t_i$, detta “tempo di campionamento” (sampling time). Viene definita la frequenza di campionamento (sampling frequency) come

$$(2.4) \quad \nu_s = \frac{1}{\Delta t}$$

Nel seguito, considerando segnali discreti, faremo riferimento implicitamente sempre a segnali campionati uniformemente.

¹ Si fa differenza tra un **tempo assoluto** e un **tempo relativo**, ovvero una **differenza temporale**. Un segnale può essere osservato a un preciso tempo (tempo assoluto), ma in genere viene descritto in tempo relativo, per esempio prendendo l'origine temporale all'inizio del segnale.

² Si distingue in genere tra *tempo assoluto* (indicata in genere con t) e *distanza temporale* (in genere indicata con τ). In questo caso parliamo di tempo assoluto.

Tipi di segnali

Oltre alla classificazione tra segnali continui e discreti, esistono altre caratteristiche peculiari di un segnale:

- Reale o complesso, cioè a un dato istante il segnale è un numero reale o un numero complesso.
- Scalare o vettoriale, cioè se è descritto da un singolo numero ad ogni istante, o da più numeri.
- Analogico o digitale, cioè il segnale ad un dato istante può assumere qualsiasi valore reale in un dato intervallo (segnale analogico) o solo un numero discreto di valori (segnale digitale), tipicamente multipli interi di un certo valore detto “quanto di conversione analogico-digitale”: è questo il caso dei segnali acquisiti da un sistema digitale tramite un convertitore ADC. In genere i segnali digitali sono anche discreti nel tempo.
- Periodico, se esiste un valore τ tale che, per tutti i valori di t , se il segnale è continuo (τ è un numero reale)

$$(2.5) \quad x(t) = x(t + \tau)$$

e se il segnale è discreto³ (τ è un numero intero)

$$(2.6) \quad x_i = x_{i+\tau}$$

- Deterministico o casuale (o stocastico): questa classificazione riguarda il “modello” del segnale: nel primo caso se è completamente definito a priori il suo valore, nel secondo se è definito solo statisticamente. I segnali deterministici possono godere di particolari simmetrie rispetto a un certo istante, per esempio all’istante $t=0$: sono detti pari se $x(t)=x(-t)$ e dispari se $x(t)=-x(-t)$.

³ In effetti questa definizione è restrittiva: se campioniamo un segnale continuo periodico, abbiamo in genere un segnale discreto che non è periodico secondo questa definizione.

Segnali notevoli

Vediamo ora alcuni semplici segnali discreti di particolare interesse teorico e pratico. Li presenteremo insieme agli analoghi segnali continui.

- Segnale costante

Continuo

$$x(t) = c$$

Discreto

$$x_i = c$$

quando $c = 0$, diciamo che c'è assenza di segnale.

- Segnale a gradino detta anche **funzione gradino di Heaviside**

Continuo

$$\begin{cases} t < 0 \rightarrow u(t) = 0 \\ t \geq 0 \rightarrow u(t) = 1 \end{cases}$$

Discreto

$$\begin{cases} i < 0 \rightarrow u_i = 0 \\ i \geq 0 \rightarrow u_i = 1 \end{cases}$$

- Segnale a delta; è la “derivata” (o la differenza nel caso discreto) del segnale a gradino. Nel caso continuo è una delta di Dirac, nel caso discreto è un segnale sempre nullo, ma che vale 1 in 0: esso è detto anche impulso unitario o funzione impulsiva discreta:

Continuo

$$\delta(t)$$

Discreto

$$\begin{cases} i = 0 \rightarrow \delta_i = 1 \\ i \neq 0 \rightarrow \delta_i = 0 \end{cases}$$

e quindi

$$u(t) = \int_{-\infty}^t \delta(\kappa) \cdot d\kappa$$

$$u_i = \sum_{k=-\infty}^i \delta_k$$

- Segnale sinusoidale

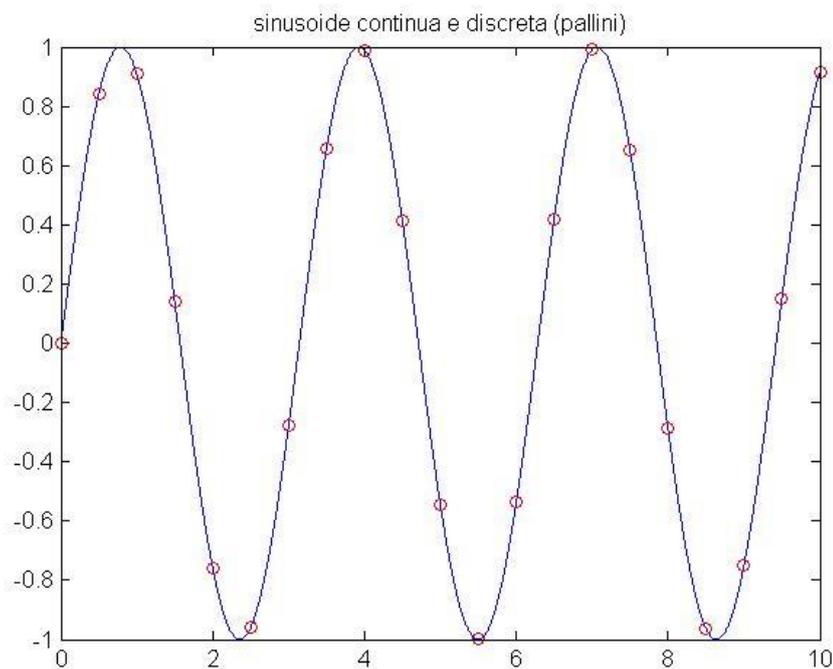
Continuo

Discreto

$$x(t) = A \cdot \sin(\omega \cdot t + \varphi)$$

$$x_i = A \cdot \sin(\omega \cdot \Delta t \cdot i + \varphi)$$

Nella figura il segnale continuo è indicato con la linea continua e quello discreto con i pallini:



Notare che il segnale sinusoidale discreto non è in genere periodico secondo la definizione che abbiamo dato in precedenza.

- Segnale esponenziale complesso⁴

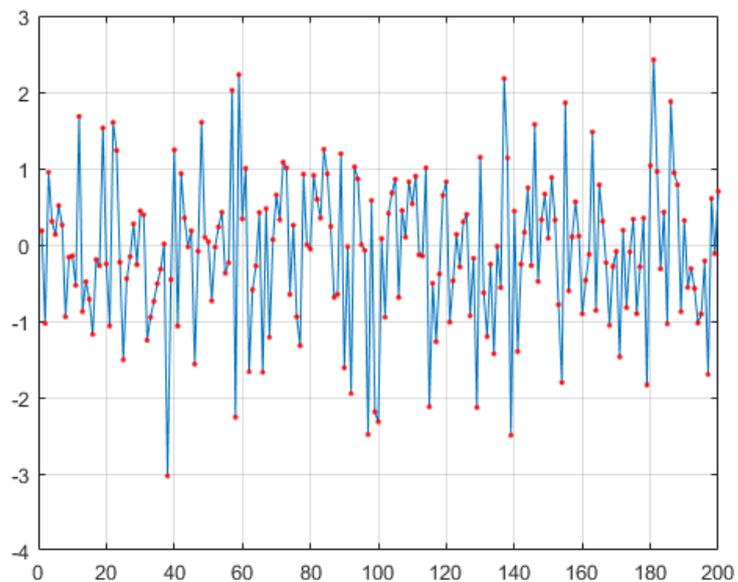
Continuo

Discreto

$$x(t) = A \cdot \exp\left[-\frac{t}{\tau} + j \cdot (\omega \cdot t + \varphi)\right]$$

$$x_i = A \cdot \exp\left[-\frac{\Delta t \cdot i}{\tau} + j \cdot (\omega \cdot \Delta t \cdot i + \varphi)\right]$$

- Rumore bianco, il più semplice segnale stocastico, se ne parlerà nel capitolo sui processi stocastici.



⁴ Indichiamo con j l'unità immaginaria.

3 - Segnali transitori

Parametri che caratterizzano un segnale transitorio

Un segnale transitorio (detto anche transiente o impulsivo) è un segnale che va a zero per il tempo tendente all'infinito in entrambi i sensi. Chiamiamo **energia del segnale**, la grandezza E_w definita nel modo seguente:

- Caso continuo, segnale $w(t)$

$$(3.1) \quad E_w = \int_{-\infty}^{\infty} |w(t)|^2 \cdot dt$$

- Caso discreto, segnale w_i

$$(3.2) \quad E_w = \sum_{i=-\infty}^{\infty} |w_i|^2$$

e supporremo inoltre che sia E_w finito.

È spesso comodo definire la “**posizione**” e la “**lunghezza**” (o “larghezza” o “durata”) di un segnale i parametri analoghi al valor medio e alla deviazione standard delle distribuzioni di probabilità :

- Caso continuo, segnale $w(t)$

$$(3.3) \quad \text{posizione} = t_w = \int_{-\infty}^{\infty} t \cdot \frac{|w(t)|^2}{E_w} \cdot dt$$

$$(3.4) \quad \text{lunghezza} = l_w = 2 \cdot \sqrt{\int_{-\infty}^{\infty} (t - t_w)^2 \cdot \frac{|w(t)|^2}{E_w} \cdot dt}$$

- Caso discreto, segnale w_i

$$(3.5) \quad \text{posizione} = t_w = \sum_{i=-\infty}^{\infty} i \cdot \frac{|w_i|^2}{E_w}$$

$$(3.6) \quad \text{lunghezza} = l_w = 2 \cdot \sqrt{\sum_{i=-\infty}^{\infty} (i - t_w)^2 \cdot \frac{|w_i|^2}{E_w}}$$

(notare il fattore 2 rispetto alla deviazione standard, che è assimilabile a una “semi-larghezza”).

Se il segnale w è reale positivo, sono spesso più comode le definizioni di posizione e durata analoghe alle precedenti, ma che utilizzano $w(t)$ o, nel caso discreto, w_i (normalizzate con il loro integrale o somma) invece che i loro moduli quadri.

Facciamo due esempi.

Prendiamo il segnale “gaussiano”

$$(3.7) \quad w(t) = A \cdot \exp\left(-\frac{t^2}{2 \cdot \sigma^2}\right)$$

Se prendiamo quest’ultima definizione di larghezza, abbiamo che la larghezza è 2σ . Con la definizione precedente abbiamo

$$|w(t)|^2 = \left| A \cdot \exp\left(-\frac{t^2}{2 \cdot \sigma^2}\right) \right|^2 = A^2 \cdot \exp\left(-\frac{t^2}{\sigma^2}\right)$$

e quindi

$$(3.8) \quad l_w = \sqrt{2} \cdot \sigma$$

cioè un valore più piccolo.

Prendiamo ora un segnale pari a 1 durante il periodo da 0 a T e nullo altrove. In questo caso le due definizioni producono lo stesso valore, ma non quello che ci piacerebbe,

cioè T , ma $\frac{T}{\sqrt{3}}$. Come è chiaro, non esiste una definizione completamente

soddisfacente di larghezza. Una possibilità ragionevole, se si vuole intendere un parametro simile alla lunghezza di un segnale costante, è, nel caso discreto,

$$(3.9) \quad \text{lunghezza} = l_w = \sqrt{6 \cdot \sum_{i=-\infty}^{\infty} (i - t_w)^2 \cdot \frac{|w_i|^2}{E_w}}$$

Parametri che caratterizzano due impulsi

Dati due impulsi x_i e y_i , possiamo calcolare alcune utili parametri e funzioni⁵:

- il **prodotto scalare**, detto anche **cross-energia**

$$(3.10) \quad r_{xy} = \sum_{i=-\infty}^{\infty} x_i \cdot y_i^*$$

- la **distanza** (teorema di Pitagora)

$$(3.11) \quad d_{xy} = \sqrt{\sum_{i=-\infty}^{\infty} |x_i - y_i|^2}$$

- la **convoluzione**,

$$(3.12) \quad z_k = \sum_{i=-\infty}^{\infty} x_{k-i} \cdot y_i$$

Vedremo che questa operazione è collegata al prodotto di due polinomi (finiti o infiniti).

- la **cross-correlazione** o **correlazione incrociata**

$$(3.13) \quad R_{xy}(k) = \sum_{i=-\infty}^{\infty} x_{k+i} \cdot y_i^*$$

Notare le differenze di definizione tra convoluzione e cross-correlazione.

⁵ Qui vengono date le espressioni per dati discreti. Analoghi concetti sono sviluppati nel caso continuo.

Trasformata di Fourier di un segnale ad energia finita

(Vedremo nel prossimo capitolo una più dettagliata presentazione della trasformata di Fourier discreta).

Un segnale transitorio, o impulso di energia finita, può essere Fourier-trasformato

$$(3.14) \quad W(\omega) = \int_{-\infty}^{\infty} w(t) \cdot e^{-j\omega t} \cdot dt$$

Vale il teorema di Parseval per cui

$$(3.15) \quad \int_{-\infty}^{\infty} |w(t)|^2 \cdot dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |W(\omega)|^2 \cdot d\omega = \int_{-\infty}^{\infty} |W(2\pi \cdot \nu)|^2 \cdot d\nu$$

dove $\nu = \frac{\omega}{2\pi}$ è la frequenza.

Nel discreto, usando la DTFT (la trasformata discreta di Fourier), abbiamo

$$(3.16) \quad W(\Omega) = \sum_{i=-\infty}^{\infty} w_i \cdot e^{-j\Omega i}$$

(dove $\Omega = \omega \cdot \Delta t$ è la pulsazione normalizzata).

Il teorema di Parseval nel discreto diventa

$$(3.17) \quad \sum_{i=-\infty}^{\infty} |w_i|^2 = \frac{1}{2\pi} \int_0^{2\pi} |W(\Omega)|^2 \cdot d\Omega$$

Notare il diverso range di integrazione tra questa e la (3.15).⁶

Analogamente a come abbiamo fatto per $w(t)$, possiamo definire posizione e larghezza per $W(\Omega)$.

È importante notare che più un impulso è “largo” nel dominio del tempo, tanto più è stretto nel dominio della frequenza. Quindi il prodotto delle larghezze nei due domini

⁶ Notare che nel discreto esiste un’unità di tempo “preferenziale”: il tempo di campionamento Δt .

non può essere inferiore a un dato valore, che dipende da come viene calcolata la larghezza e dalla forma dell'impulso nel tempo.

Questo fenomeno è detto “**principio di indeterminazione di Fourier**”, il che è comprensibile se si considera la larghezza di un impulso come l'incertezza sulla posizione, sia nel dominio del tempo che della frequenza.

Per esempio, nel caso di un impulso gaussiano con un deviazione standard σ , la trasformata di Fourier ha una forma gaussiana, con deviazione standard $1/\sigma$.

Attenzione ! Se il segnale è reale, la trasformata di Fourier è una funzione hermitiana e quindi $|W(\omega)|^2$, che è usata per calcolare la larghezza, è simmetrica rispetto all'origine. Questo fa sì che, nel caso di $w(t)$ “a banda stretta”, la larghezza di $W(\omega)$ sia molto più grande di quella che sarebbe se si considerassero solo le frequenze positive.

Autocorrelazione e spettro di energia

Se facciamo la correlazione incrociata di un segnale con se stesso, abbiamo l'autocorrelazione

$$(3.18) \quad R_x(k) = \sum_{i=-\infty}^{\infty} x_{k+i} \cdot x_i^*$$

Questa definizione di autocorrelazione per i segnali impulsivi è ovviamente diversa, ma analoga, dall'autocorrelazione nel caso dei processi stocastici, che discuteremo nel capitolo sui segnali aleatori. Come l'autocorrelazione nel caso dei processi stocastici, gode di varie proprietà:

- è la convoluzione del segnale con se stesso invertito temporalmente (e coniugato, se il segnale è complesso)
- è simmetrica rispetto a 0
- ha ivi il suo massimo assoluto (e in questo caso in 0 vale 1)
- la sua trasformata di Fourier è sempre positiva (ed è chiamato **spettro di energia**)
- la "larghezza" dell'autocorrelazione (definita come in (3.6) e (3.4)) è proporzionale a quella del segnale.

Possiamo definire spettro di energia la trasformata di Fourier dell'autocorrelazione

$$(3.19) \quad S_x(\omega) = \sum_k R_x(k) \cdot e^{-j\omega k}$$

essendo $R_x(k)$ ricavato dalla (3.18).

Vale il seguente risultato (teorema di Wiener-Kinchin):

$$(3.20) \quad S_x(\omega) = \left| \sum_k x_k \cdot e^{-j\omega k} \right|^2$$

Quindi possiamo calcolare lo spettro di energia senza calcolare l'autocorrelazione.

4 - Trasformate

Live Script [sfft.mlx](#), [Hilbert.mlx](#)

Trasformata z

Analogia con la trasformata di Laplace

Un segnale discreto $\{x_i\}$ può essere associato ad un segnale continuo, costituito da una successione di delte di Dirac,

$$(4.1) \quad x(t) = \sum_k x_k \cdot \delta(t - k \cdot \Delta t)$$

essendo Δt il tempo di campionamento. La trasformata di Laplace del segnale $x(t)$ è

$$(4.2) \quad X_L(s) = \sum_k x_k \cdot e^{-s \cdot k \cdot \Delta t}$$

e la trasformata di Fourier è

$$(4.3) \quad X_F(\omega) = \sum_k x_k \cdot e^{-j\omega k \cdot \Delta t}$$

Ricordiamo che nella trasformata di Laplace $e^{-s \cdot \Delta t}$ equivale a un operatore ritardo di Δt (come s è l'operatore derivata).

Sostituendo $z = e^{s \cdot \Delta t}$, si ha

$$(4.4) \quad X_L(s) = X(z) \Big|_{z=e^{s\Delta t}}$$

e, se sul cerchio di raggio 1 c'è convergenza,

$$(4.5) \quad X_F(\omega) = X(z) \Big|_{z=e^{j\omega\Delta t}}$$

Si noti che $X_L(s)$ e $X_F(\omega)$ sono periodiche (la prima sulla direzione immaginaria), con periodo

$$(4.6) \quad \omega_0 = \frac{2\pi}{\Delta t}$$

quindi possono essere definite per un solo periodo, nel range di frequenza da 0 a $\frac{1}{\Delta t}$ o da $-\frac{1}{2 \cdot \Delta t}$ a $\frac{1}{2 \cdot \Delta t}$.

Il mapping del piano s sul piano z (non univoco, il che spiega la suddetta periodicità) è tale che l'asse immaginario della trasformata di Laplace si trasforma nel cerchio unitario della trasformata z, il semi-piano sinistro diventa l'interno di tale cerchio e il semi-piano destro l'esterno.

La trasformata z di una successione di campioni equispaziati $\{x_k\} = \{x_1 \ x_2 \ \dots \ x_n\}$ è

$$(4.7) \quad X(z) = \sum_{k=1}^n x_k \cdot z^{-k}$$

In generale l'indice k può andare da $-\infty$ a ∞ , eventualmente con valori nulli dei campioni per gli indici non esistenti

$$(4.8) \quad X(z) = \sum_{k=-\infty}^{\infty} x_k \cdot z^{-k}$$

Quindi la trasformata z trasforma una successione di numeri (in genere reali o complessi) in un polinomio.

Si noti che, se la serie $\{x_i\}$ è di lunghezza finita, la regione di convergenza copre tutto il piano della variabile z (piano di Gauss). Se è *infinita destra*, cioè nulla per i minore di un certo valore, la regione di convergenza è l'esterno di un cerchio di centro l'origine (il raggio del cerchio dipende dalla serie). Analogamente, se è *infinita sinistra*, cioè nulla per i maggiore di un certo valore, la regione di convergenza è l'interno di un cerchio di centro l'origine.

Se la trasformata X(z) converge nel dominio anulare definito da $R_1 < |z| < R_2$ allora si può calcolare la trasformata inversa con

$$(4.9) \quad x_k = \frac{1}{2\pi j} \oint_C X(z) \cdot z^{k-1} \cdot dz$$

dove C è un percorso chiuso che separa $|z|=R_1$ da $|z|=R_2$. Questo integrale può calcolarsi con l'ausilio del teorema dei residui di Cauchy:

$$(4.10) \quad \frac{1}{2\pi j} \oint_C X(z) \cdot z^{k-1} \cdot dz = \sum [\text{residui di } X(z) \cdot z^{k-1} \text{ nei poli interni a } C]$$

Se la $X(z)$ è una funzione razionale, se cioè è data dal rapporto di due polinomi, possiamo ottenere il rapporto dei due polinomi come un polinomiale della variabile z^{-1} , l'antitrasformata si ottiene immediatamente dai coefficienti delle varie potenze di z .

In molti casi la trasformata inversa può calcolarsi usando le tavole e le proprietà della trasformata z . A volte, nel caso di trasformate che sono funzioni razionali, basta eseguire la divisione tra il polinomio numeratore e il polinomio divisore e considerare i coefficienti del polinomio risultato; per esempio se

$$(4.11) \quad X(z) = \frac{1}{1-w \cdot z^{-1}}$$

con w un numero complesso; dividendo 1 per $(1-wz^{-1})$, si ha

$$(4.12) \quad X(z) = \frac{1}{1-w \cdot z^{-1}} = 1 + w \cdot z^{-1} + w^2 \cdot z^{-2} + w^3 \cdot z^{-3} + \dots$$

e quindi l'antitrasformata è $\{x_k\} = \{1 \ w \ w^2 \ w^3 \ \dots\}$.

Proprietà della trasformata z

Sia $X(z)$, $X^{(1)}(z)$ e $X^{(2)}(z)$ le trasformate di $\{x_i\}$, $\{x_i^{(1)}\}$ e $\{x_i^{(2)}\}$ rispettivamente.

- **Linearità:**

$$(4.13) \quad \{a \cdot x_i^{(1)} + b \cdot x_i^{(2)}\} \leftrightarrow \{a \cdot X^{(1)}(z) + b \cdot X^{(2)}(z)\}$$

- **Shift (scorrimento, ovvero ritardo o avanzamento):**

$$(4.14) \quad \{x_{i-k}\} \leftrightarrow z^{-k} \cdot X(z)$$

Notiamo che la variabile z agisce come un operatore di avanzamento, ovvero uno scorrimento a sinistra; z^{-1} come un operatore di ritardo, ovvero uno scorrimento a destra.

Nelle successioni discrete l'operatore differenza ha una funzione simile all'operatore derivata per le funzioni continue. La differenza prima $\{x_k - x_{k-1}\}$ ha trasformata

$$(4.15) \quad \{x_k - x_{k-1}\} \leftrightarrow X(z) \cdot (1 - z^{-1})$$

L'operatore "differenza prima" è quindi $1 - z^{-1}$. L'operatore "differenza i-esima" (analoga alla derivata i-esima) è $(1 - z^{-1})^i$.

- **Inversione temporale :**

$$(4.16) \quad \{x_{-i}\} \leftrightarrow X(z^{-1})$$

- **Moltiplicazione per una sequenza esponenziale :**

$$(4.17) \quad \{\alpha^i \cdot x_i\} \leftrightarrow X\left(\frac{1}{\alpha} z\right)$$

Se α è reale, ciò corrisponde ad uno scaling nel dominio z , se è $\alpha = e^{j\varphi}$, ciò corrisponde ad una rotazione nel piano z .

- **Convoluzione :**

La trasformata z del segnale ottenuto come convoluzione di due segnali discreti (vedi equazione (3.12)) è data dal prodotto delle due trasformate z dei due segnali. Essendo x e y due successioni con trasformata z rispettivamente X e Y, si ha

$$(4.18) \quad x \otimes y \leftrightarrow X \cdot Y$$

La dimostrazione di questa proprietà, nel caso di segnali discreti di lunghezza finita, è abbastanza semplice: è legata all'**analogia tra l'operazione di convoluzione discreta e il prodotto di due polinomi**. Per rendersene conto, sviluppare il prodotto

$$(4.19) \quad \begin{aligned} & (x_0 + x_1 \cdot a + x_2 \cdot a^2 + \dots + x_n \cdot a^n) \cdot (y_0 + y_1 \cdot a + y_2 \cdot a^2 + \dots + y_m \cdot a^m) = \\ & (x_0 \cdot y_0) + (x_0 \cdot y_1 + x_1 \cdot y_0) \cdot a + (x_0 \cdot y_2 + x_1 \cdot y_1 + x_2 \cdot y_0) \cdot a^2 + \dots + (x_0 \cdot y_m) \cdot a^{m+n} \end{aligned}$$

- **Coniugazione :**

$$(4.20) \quad \{x_i^*\} \leftrightarrow X^*(z^*)$$

Da cui si deduce che, se x è reale,

$$(4.21) \quad X(z) = X^*(z^*)$$

Alcune trasformate z

Vediamo ora le trasformate z di alcune semplici successioni:

- delta discreta

$$(4.22) \quad \delta_i = \begin{cases} 0 & \text{per } i \neq 0 \\ 1 & \text{per } i = 0 \end{cases} \leftrightarrow 1$$

- funzione gradino

$$(4.23) \quad u_i = \begin{cases} 0 & \text{per } i < 0 \\ 1 & \text{per } i \geq 0 \end{cases} \leftrightarrow \sum_{i=0}^{\infty} z^{-i} = \frac{1}{1-z^{-1}} = \frac{z}{z-1}$$

- esponenziale

$$(4.24) \quad u_i w^i \leftrightarrow \frac{1}{1-w \cdot z^{-1}} = \frac{z}{z-w}$$

- esponenziale "rampato"

$$(4.25) \quad u_i \cdot i \cdot w^i \leftrightarrow \frac{w \cdot z^{-1}}{(1-w \cdot z^{-1})^2} = \frac{wz}{(z-w)^2}$$

- coseno

$$(4.26) \quad u_i \cdot \cos(\omega_0 \cdot i) \leftrightarrow \frac{1 - \cos(\omega_0) \cdot z^{-1}}{1 - 2 \cos(\omega_0) \cdot z^{-1} + z^{-2}} = \frac{z^2 - \cos(\omega_0) \cdot z}{z^2 - 2 \cos(\omega_0) \cdot z + 1}$$

- seno

$$(4.27) \quad u_i \cdot \sin(\omega_0 \cdot i) \leftrightarrow \frac{\sin(\omega_0) \cdot z^{-1}}{1 - 2 \cos(\omega_0) \cdot z^{-1} + z^{-2}} = \frac{\sin(\omega_0) \cdot z}{z^2 - 2 \cos(\omega_0) \cdot z + 1}$$

Trasformata di Fourier

Ricordiamo le principali proprietà della trasformata di Fourier nel tempo continuo.

Partiamo dallo sviluppo in serie di Fourier. Se una funzione è periodica, cioè se $x(t) = x(t + T_0)$ per tutti i t (T_0 è il periodo, $\nu_0 = \frac{1}{T_0}$ è la frequenza e $\omega_0 = 2\pi\nu_0$ la pulsazione), come è noto possiamo svilupparla in serie di Fourier, cioè

$$(4.28) \quad x(t) = \sum_{k=-\infty}^{\infty} X_k \cdot e^{jk\omega_0 t}$$

dove i coefficienti X_k sono ricavati da

$$(4.29) \quad X_k = \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} x(t) \cdot e^{-j \cdot k \cdot \frac{2\pi}{T_0} \cdot t} \cdot dt$$

La trasformata di Fourier è una generalizzazione dello sviluppo in serie di Fourier al caso in cui la funzione $x(t)$ non sia periodica (ovvero sia di periodo infinito).

La trasformata di Fourier è simile alla trasformata di Laplace; in essa però la variabile “trasformata”, coniugata del tempo, è reale e indica la pulsazione (in Inglese *angular frequency*; spesso viene indicata con ω). Ricordiamo che $\omega = 2\pi \cdot \nu$, dove ν è la frequenza.

La trasformata di Fourier di una funzione $x(t)$ è data da⁷

$$(4.30) \quad X(\omega) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j\omega t} dt$$

Come si vede è un caso particolare della trasformata (bilatera) di Laplace, dove $s = j\omega$. Le funzioni della base trasformata sono, come si vede, esponenziali complesse (ricordiamo che $e^{j\omega t} = \cos(\omega \cdot t) + j \cdot \sin(\omega \cdot t)$).

Nella trasformata di Fourier c'è una simmetria perfetta tra la variabile t e la variabile ω . In pratica però, molto spesso nel dominio t le funzioni sono reali mentre sono complesse (ed hermitiane) nel dominio ω .

Notiamo che

$$(4.31) \quad X(0) = \int_{-\infty}^{\infty} x(t) dt$$

⁷ Attenzione ! Esistono altre definizioni, quasi equivalenti; per esempio con l'esponenziale positivo, oppure con costanti a moltiplicare differenti.

La trasformata inversa si ottiene da

$$(4.32) \quad x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) \cdot e^{j\omega t} d\omega$$

Essa viene anche indicata con $X(\omega) = \mathcal{F}\{x(t)\}$ o, quando è chiaro che parliamo di trasformate di Fourier, con $x(t) \leftrightarrow X(\omega)$. Se $X_L(s)$ è la trasformata di Laplace di $x(t)$ e $X_F(\omega)$ la trasformata di Fourier, allora⁸ $X_F(\omega) = X_L(j\omega)$.

Data la parentela con la trasformata di Laplace, la trasformata di Fourier gode di proprietà simili, più altre dovute alla maggiore "simmetria" tra le variabili coniugate.

Guardiamo con attenzione queste proprietà, poiché nell'analisi dei segnali si fa largo uso di questa trasformata.

Sia $x(t) \leftrightarrow X(\omega)$, $x_1(t) \leftrightarrow X_1(\omega)$, $x_2(t) \leftrightarrow X_2(\omega)$. Si ha:

- **Linearità:** $a \cdot x_1(t) + b \cdot x_2(t) \leftrightarrow a \cdot X_1(\omega) + b \cdot X_2(\omega)$

Se conosciamo la trasformata di due segnali, possiamo calcolare immediatamente quella di una combinazione di questi due segnali. Possiamo, quando è il caso, scomporre opportunamente un segnale per calcolare, o intuire, la sua trasformata.

- **Spostamento nel tempo:** $x(t-t_0) \leftrightarrow e^{-j\omega t_0} X(\omega)$

Se trasliamo un segnale nel tempo, per la trasformata abbiamo lo stesso valore assoluto e uno sfasamento che varia linearmente con ω .

Si noti che $|X(\omega)|$ è indipendente da spostamenti nel tempo.

- **Spostamento nel dominio ω :** $e^{j\omega_0 t} x(t) \leftrightarrow X(\omega - \omega_0)$

Questa è la duale della precedente. $e^{j\omega_0 t}$ rappresenta una oscillazione complessa: moltiplicare $x(t)$ per questa oscillazione complessa provoca una traslazione della trasformata di Fourier.

Se $x(t)=1$, troviamo che

$$(4.33) \quad e^{j\omega_0 t} \leftrightarrow \delta(\omega - \omega_0)$$

⁸ Nel caso in cui esistono entrambe.

Ora notiamo che $e^{j\omega_0 t} + e^{-j\omega_0 t} = 2 \cdot \cos(\omega_0 t)$ e $e^{j\omega_0 t} - e^{-j\omega_0 t} = 2j \cdot \sin(\omega_0 t)$ allora

$$(4.34) \quad x(t) \cdot \cos(\omega_0 t) \leftrightarrow \frac{X(\omega - \omega_0) + X(\omega + \omega_0)}{2}$$

$$(4.35) \quad x(t) \cdot \sin(\omega_0 t) \leftrightarrow \frac{X(\omega - \omega_0) - X(\omega + \omega_0)}{2j}$$

- **Cambiamento di scala temporale:** $x(at) \leftrightarrow \frac{1}{|a|} X\left(\frac{\omega}{a}\right)$

Anche questa è una proprietà molto importante. la trasformazione $x'(t) = x(a \cdot t)$ produce una contrazione o una dilatazione a seconda che $|a|$ sia maggiore o minore di 1. Una dilatazione nel dominio t corrisponde a una contrazione nel dominio ω e viceversa.

Questa proprietà è la ragione del cosiddetto “principio di indeterminazione di Fourier”, simile a quello di Heisenberg, che vedremo in seguito.

- **Inversione temporale:** $x(-t) \leftrightarrow X(-\omega)$

Se la $x(t)$ è reale, allora $X(\omega) = X^*(-\omega)$ e quindi anche $x(-t) \leftrightarrow X^*(\omega)$

- **Differenziazione nel dominio del tempo:** $\frac{d x(t)}{d t} \leftrightarrow j\omega \cdot X(\omega)$

Questa proprietà può essere generalizzata con

$$(4.36) \quad \frac{d^n x(t)}{d t^n} \leftrightarrow (j\omega)^n \cdot X(\omega)$$

- **Integrazione nel dominio del tempo:** $\int_{-\infty}^t x(\tau) d\tau \leftrightarrow \pi X(0)\delta(\omega) + \frac{1}{j\omega} X(\omega)$

- **Differenziazione nel dominio ω :** $-j \cdot t \cdot x(t) \leftrightarrow \frac{dX(\omega)}{d\omega}$

- **Convoluzione:** $\int_{-\infty}^{\infty} x_1(\tau) x_2(t-\tau) d\tau \leftrightarrow X_1(\omega) \cdot X_2(\omega)$

Sostituire una operazione come la convoluzione, spesso complicata, se eseguita analiticamente, o computazionalmente costosa, se eseguita numericamente, con una moltiplicazione rende spesso molto utile lavorare nel dominio trasformato di Fourier.

- **Moltiplicazione:** $x_1(t) \cdot x_2(t) \leftrightarrow \int_{-\infty}^{\infty} X_1(\theta) X_2(\omega-\theta) d\theta$

È la duale della proprietà della convoluzione.

- **Identità di Parseval:**

$$\int_{-\infty}^{\infty} x_1(t) \cdot x_2^*(t) \cdot dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_1(\omega) \cdot X_2^*(\omega) \cdot d\omega$$

$$\int_{-\infty}^{\infty} |x(t)|^2 \cdot dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2 \cdot d\omega$$

Questa proprietà è un'utilissima relazione sull' "energia" totale del segnale nel dominio del tempo e di ω . Si noti che se l'integrale si esegue nella variabile ν non compare il "fastidioso" coefficiente $\frac{1}{2\pi}$.

Inoltre, se il segnale $x(t)$ è reale, la sua trasformata di Fourier è $X(\omega) = X^*(-\omega)$, cioè la sua parte reale è pari e la parte immaginaria dispari (funzione hermitiana); se $x(t)$, reale o complessa, è una funzione hermitiana, la trasformata è reale, se è anti-hermitiana ($x(t) = -x^*(-t)$), è immaginaria.

Ecco alcuni casi:

Dominio del tempo	Dominio della frequenza
Reale	Hermitiana
Reale positiva	Massimo reale in 0, hermitiana
Reale pari o hermitiana	Reale
Reale dispari o anti-hermitiana	Immaginaria

Ricordiamo qui che, se le dimensioni di $x(t)$ sono $[x]$, quelle della trasformata X sono $[xt]$.

Un risultato importante per i sistemi che si deduce dalla proprietà della trasformata di Fourier per la convoluzione, è che se l'ingresso di un sistema lineare è una senoide, l'uscita è anch'essa una senoide, con ampiezza e fase che dipendono dalla frequenza. Quindi per ogni frequenza possiamo definire un numero complesso che abbia come modulo il "guadagno" del sistema, cioè il rapporto tra l'ampiezza della senoide in uscita e quella in ingresso, e la fase la differenza di fase tra di esse. La funzione di trasferimento (definita come la trasformata di Laplace o di Fourier della risposta impulsiva; la vedremo meglio in seguito) dà, per ogni frequenza, questo numero complesso.

Per quanto riguarda i segnali, la trasformata di Fourier ci dà il contenuto in energia (o, come vedremo per i segnali di durata infinita, in potenza) alle varie frequenze. Definiamo spettro di energia di un dato segnale, il modulo quadro della sua trasformata di Fourier.

Infine riportiamo la trasformata di Fourier di alcune significative funzioni:

- **Gaussiana:**

$$(4.37) \quad \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{t^2}{2\sigma^2}} \leftrightarrow e^{-\frac{\omega^2 \cdot \sigma^2}{2}}$$

Notiamo che la trasformata ha la stessa forma funzionale della gaussiana (ma mentre nel tempo è normalizzata a 1, in ω no). Se prendiamo come "larghezza" nel tempo la deviazione standard, in ω la larghezza è $\frac{1}{\sigma}$, quindi il prodotto delle due larghezze è sempre 1.

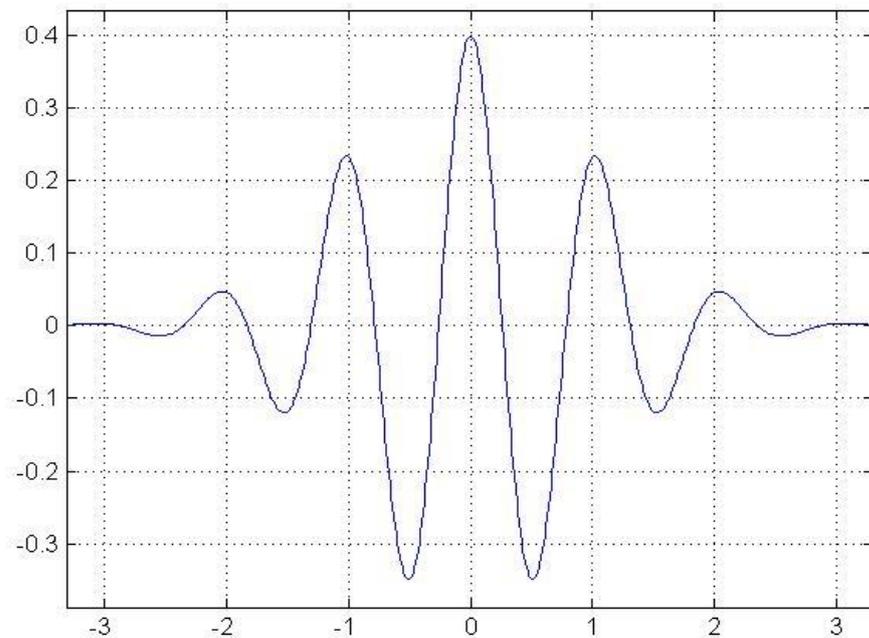
- **Pacchetto gaussiano:**

Un “pacchetto” è un segnale sinusoidale moltiplicato per una “finestra”. In questo caso la finestra è gaussiana. Il calcolo della trasformata è facile, ricordando le proprietà su riportate. Sono qui riportati i tre casi del pacchetto esponenziale complesso (dati complessi, un solo picco nella trasformata), del pacchetto cosinusoidale (funzione reale pari, trasformata reale) e del pacchetto sinusoidale (funzione dispari, trasformata immaginaria).

Notiamo poi che più è “largo” il pacchetto nel tempo, più sono stretti i picchi (o il picco, nel caso complesso) nella trasformata.

$$\begin{aligned}
 & \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{t^2}{2\sigma^2}} \cdot e^{j\omega_0 t} \leftrightarrow e^{-\frac{(\omega-\omega_0)^2 \cdot \sigma^2}{2}} \\
 (4.38) \quad & \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{t^2}{2\sigma^2}} \cdot \cos(\omega_0 t) \leftrightarrow \frac{e^{-\frac{(\omega-\omega_0)^2 \cdot \sigma^2}{2}} + e^{-\frac{(\omega+\omega_0)^2 \cdot \sigma^2}{2}}}{2} \\
 & \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{t^2}{2\sigma^2}} \cdot \sin(\omega_0 t) \leftrightarrow j \cdot \frac{e^{-\frac{(\omega+\omega_0)^2 \cdot \sigma^2}{2}} - e^{-\frac{(\omega-\omega_0)^2 \cdot \sigma^2}{2}}}{2}
 \end{aligned}$$

In figura c'è un pacchetto gaussiano cosinusoidale



- **Rettagolo:**

$$(4.39) \quad r(t;a) = \begin{cases} 1 & \text{per } |t| < a \\ 0 & \text{per } |t| > a \end{cases} \leftrightarrow 2a \cdot \frac{\sin(a \cdot \omega)}{a \cdot \omega}$$

Se prendiamo come misura della larghezza delle due funzioni il valore più basso in cui si azzerano, abbiamo che nel dominio t è a , nel dominio ω è $\frac{\pi}{a}$.

È stata introdotta la funzione sinc(x) come

$$(4.40) \quad \text{sinc}(x) = \frac{\sin(x)}{x}$$

e la trasformata dell'impulso rettangolare può scriversi in termini di questa funzione. La sinc ha una notevole importanza nello studio del campionamento di segnali continui.

- **Pacchetto rettangolare:**

La situazione è analoga al caso del pacchetto gaussiano.

...

- **Esponenziale simmetrico \leftrightarrow Lorentziana (distribuzione di Cauchy):**

$$(4.41) \quad e^{-\frac{|x|}{\tau}} \leftrightarrow \frac{\frac{2}{\tau}}{x^2 + \left(\frac{1}{\tau}\right)^2}$$

- **Pacchetto esponenziale simmetrico:**

La situazione è analoga al caso del pacchetto gaussiano.

...

- **delta:**

$$(4.42) \quad \delta(t) \leftrightarrow 1$$

- **costante:**

$$(4.43) \quad x(t) = c \leftrightarrow 2\pi c \cdot \delta(0)$$

- **gradino:**

$$(4.44) \quad u(t) \leftrightarrow \frac{1}{j\omega} + \pi \cdot \delta(\omega)$$

- esponenziale complesso:

$$(4.45) \quad e^{j\omega_0 t} \leftrightarrow 2\pi \cdot \delta(\omega - \omega_0)$$

- esponenziale decrescente:

$$(4.46) \quad u(t) \cdot e^{-\frac{t}{\tau}} \leftrightarrow \frac{1}{\frac{1}{\tau} + j\omega}$$

Si noti che la stessa funzione invertita nel tempo ha trasformata $\frac{1}{\frac{1}{\tau} - j\omega}$. Sommando

la $u(t) \cdot e^{-\frac{t}{\tau}}$ e la invertita nel tempo, si ha una funzione che ha come trasformata la somma delle trasformate

$$(4.47) \quad u(t) \cdot e^{-\frac{t}{\tau}} + u(-t) \cdot e^{\frac{t}{\tau}} = e^{-\frac{|t|}{\tau}} \leftrightarrow \frac{1}{\frac{1}{\tau} + j\omega} + \frac{1}{\frac{1}{\tau} - j\omega} = \frac{\frac{2}{\tau}}{\frac{1}{\tau^2} + \omega^2}$$

che non è altro che la (4.41).

- “doppio polo”:

$$(4.48) \quad u(t) \cdot t \cdot e^{-\frac{t}{\tau}} \leftrightarrow \frac{1}{\left(\frac{1}{\tau} + j\omega\right)^2}$$

Ricordiamo che la trasformata di Fourier (continua) può essere vista come la trasformata di Laplace calcolata sull'asse immaginario ($j\omega$). La trasformata di Fourier discreta può essere vista come la trasformata z calcolata sul cerchio unitario ($e^{j\omega}$).

La trasformata di Fourier per dati discreti (DTFT)

Come per i segnali continui abbiamo fatto derivare la trasformata di Fourier dalla trasformata di Laplace, così per i segnali discreti possiamo derivare dalla trasformata z la trasformata di Fourier per dati discreti (DTFT, discrete time Fourier transform) . Dalla (4.8), imponendo $|z|=1$, cioè $z = e^{j\Omega}$, dove

$$(4.49) \quad \Omega = \omega \cdot \Delta t$$

è la pulsazione normalizzata (essendo Δt il tempo di campionamento), a cui corrisponde, dividendola per 2π , una frequenza normalizzata a-dimensionale che ha valore massimo 1, abbiamo

$$(4.50) \quad \mathcal{X}(\Omega) = \sum_{i=-\infty}^{\infty} x_i \cdot e^{-j \cdot i \cdot \Omega}$$

La trasformata inversa si ottiene come

$$(4.51) \quad x_i = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{X}(\Omega) \cdot e^{j \cdot i \cdot \Omega} \cdot d\Omega$$

Ovviamente la (4.50) ha senso solo se la sommatoria converge e condizione sufficiente perché ciò avvenga è che $\sum |x_i| < \infty$.

Notiamo che la (4.50) è equivalente alla trasformata di Fourier per segnali continui per il segnale continuo

$$(4.52) \quad x(t) = \sum x_i \cdot \delta(t - i \cdot \Delta t)$$

Questa trasformata, oltre alle analoghe proprietà della trasformata di Fourier normale, gode delle seguenti:

- **periodicità:** la $\mathcal{X}(\Omega)$ è periodica di periodo 2π , come è ovvio

- **prima differenza** (analoga alla derivata nel continuo)

$$(4.53) \quad x_i - x_{i-1} \leftrightarrow (1 - e^{-j\Omega}) \mathcal{X}(\Omega)$$

- **accumulazione** (analogo nel discreto dell'integrazione)

$$(4.54) \quad \sum_{k=-\infty}^i x_k \leftrightarrow \pi \cdot \mathcal{X}(0) \delta(\Omega) + \frac{1}{1 - e^{-j\Omega}} \mathcal{X}(\Omega)$$

La DTFT è analoga alla serie di Fourier (vedi (4.28)), solo che a domini invertiti: è discreta nel tempo e continua e periodica nelle frequenze, come la serie di Fourier è continua e periodica nel tempo e discreta nelle frequenze.

La DFT (e la FFT)

Introduciamo ora la trasformata discreta di Fourier, un argomento di grande interesse pratico, poiché è ciò che viene in effetti normalmente calcolato col computer. Per far ciò poniamo che la $\{x_i\}$ non sia di lunghezza infinita, ma lunga N (ovvero che sia nulla per $i < 0$ e per $i > N-1$).

Definiamo trasformata discreta di Fourier (DFT) la seguente successione:

$$(4.55) \quad X_k = \sum_{i=0}^{N-1} x_i \cdot W_N^{k \cdot i}$$

dove

$$(4.56) \quad W_N = e^{-j \frac{2\pi}{N}}$$

è la radice N -esima dell'unità. Possiamo calcolare la DFT inversa come

$$(4.57) \quad x_i = \frac{1}{N} \sum_{k=0}^{N-1} X_k \cdot W_N^{-k \cdot i}$$

Come si vede, abbiamo una trasformata tra due domini discreti e finiti: infatti sia il tempo che la frequenza (più precisamente la pulsazione) sono discreti e finiti. Quindi si capisce l'utilità di questa trasformata, ovviamente parente stretto della trasformata di Fourier per i dati discreti. Infatti

$$(4.58) \quad X_k = \mathcal{X} \left(\frac{k \cdot 2\pi}{N} \right)$$

Il calcolo della DFT, per grandi valori di N (che può essere per esempio dell'ordine di 10^6), richiede un grandissima potenza di calcolo. Infatti, trascurando il calcolo delle potenze di W_N , che possono essere calcolati a priori, dobbiamo eseguire circa N^2 moltiplicazioni ed altrettante addizioni (per $N=10^6$ abbiamo 10^{12} operazioni di ciascun tipo). È stato però introdotto (da Cooley e Tukey nel 1965) un algoritmo di calcolo, la Fast Fourier Transform o FFT, che riduce il numero di operazioni a un valore proporzionale a $N \cdot \log N$ (in una delle migliori implementazioni è $5 \cdot N \cdot \log_2 N$, per $N=10^6$ abbiamo circa 10^8 , cioè 10000 volte più veloce).

Trasformata di Hilbert e segnale analitico

La trasformata di Hilbert di un segnale $x(t)$ è definita come

$$(4.59) \quad x'(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} \cdot d\tau$$

e l'antitrasformata è

$$(4.60) \quad x(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x'(\tau)}{t-\tau} \cdot d\tau$$

Questa trasformata è in effetti un filtro lineare che ha la proprietà fondamentale che

$$(4.61) \quad \begin{aligned} \sin(\omega \cdot t) &\leftrightarrow -\cos(\omega \cdot t) \\ \cos(\omega \cdot t) &\leftrightarrow \sin(\omega \cdot t) \end{aligned}$$

per questo motivo è detta anche filtro di quadratura, che “sfasa” ogni componente sinusoidale di 90 gradi. Tale filtro ha la funzione di trasferimento

$$(4.62) \quad H(j\omega) = -j \cdot \text{sign}(\omega) = \begin{cases} -j & \text{per } \omega > 0 \\ j & \text{per } \omega < 0 \end{cases}$$

Se $x'(t)$ è la trasformata di Hilbert del segnale reale $x(t)$, allora definiamo segnale analitico il segnale

$$(4.63) \quad x_A(t) = x(t) + j \cdot x'(t)$$

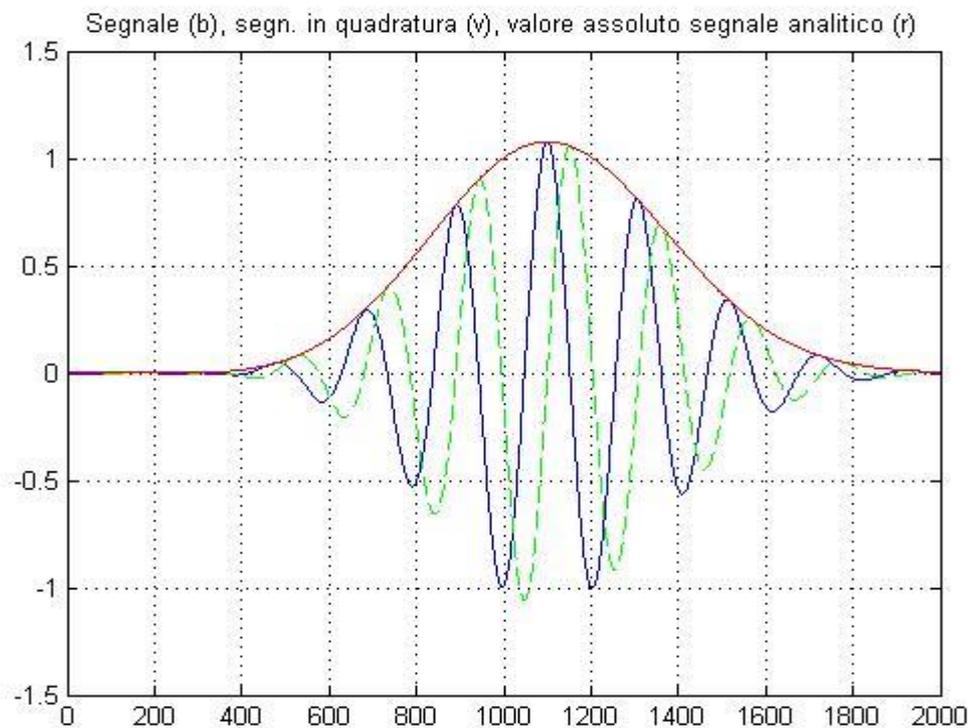
Se $X(\omega)$ è la trasformata di Fourier di $x(t)$, allora la trasformata di Fourier di $x_A(t)$ è

$$(4.64) \quad X_A(\omega) = \begin{cases} 0 & \text{per } \omega < 0 \\ 2 \cdot X(\omega) & \text{per } \omega > 0 \end{cases}$$

La trasformata di Hilbert si può definire anche nel discreto come un filtro di risposta impulsiva

$$(4.65) \quad h_i = \begin{cases} 0 & \text{per } i \text{ pari} \\ \frac{2}{\pi \cdot i} & \text{per } i \text{ dispari} \end{cases}$$

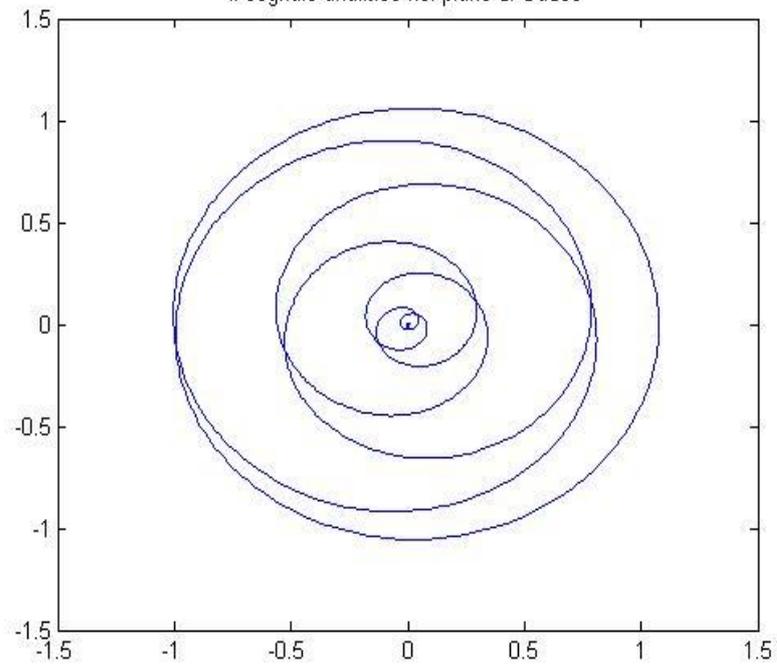
Vediamo un esempio.



Come si vede, il valore assoluto del segnale analitico è l'involuppo del segnale originario.

Ed ecco questo segnale analitico nel piano di Gauss

Il segnale analitico nel piano di Gauss



Altre trasformate

Nell'analisi dei segnali spesso vengono utilizzati altri tipi di trasformate. Queste sono per lo più collegate alla trasformata di Fourier, introducendo però una "localizzazione" nel tempo. Esempi di questo tipo sono:

- la trasformata di Gabor
- la short-time Fourier transform (STFT)
- la trasformata di Hermite
- la chirplet transform
- trasformata wavelet
- ...

La trasformata di Walsh-Hadamard utilizza un insieme di funzioni a due solo valori. In questo caso il concetto di frequenza viene sostituito da quello di "sequenza".

Altri tipi di trasformate, utilizzate spesso nell'analisi di immagini sono

- la trasformata di Radon
- la trasformata di Hough

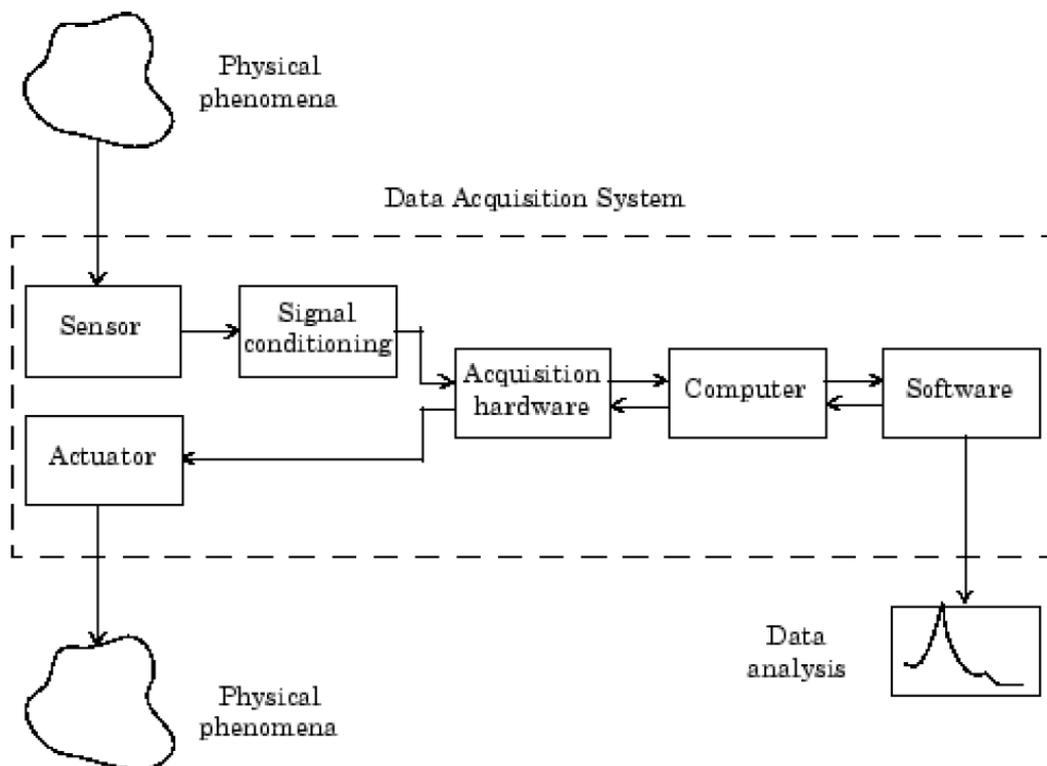
5 - Sistemi di acquisizione

[Live Script Conditioning.mlx](#), [LabAliasing.mlx](#), [Dist_Arm.mlx](#)

Vedremo ora come funziona un sistema di acquisizione dati e quali sono i problemi che si incontrano nella realizzazione.

Schema di un sistema di acquisizione

Questo è lo schema generale di un sistema di acquisizione dati⁹:



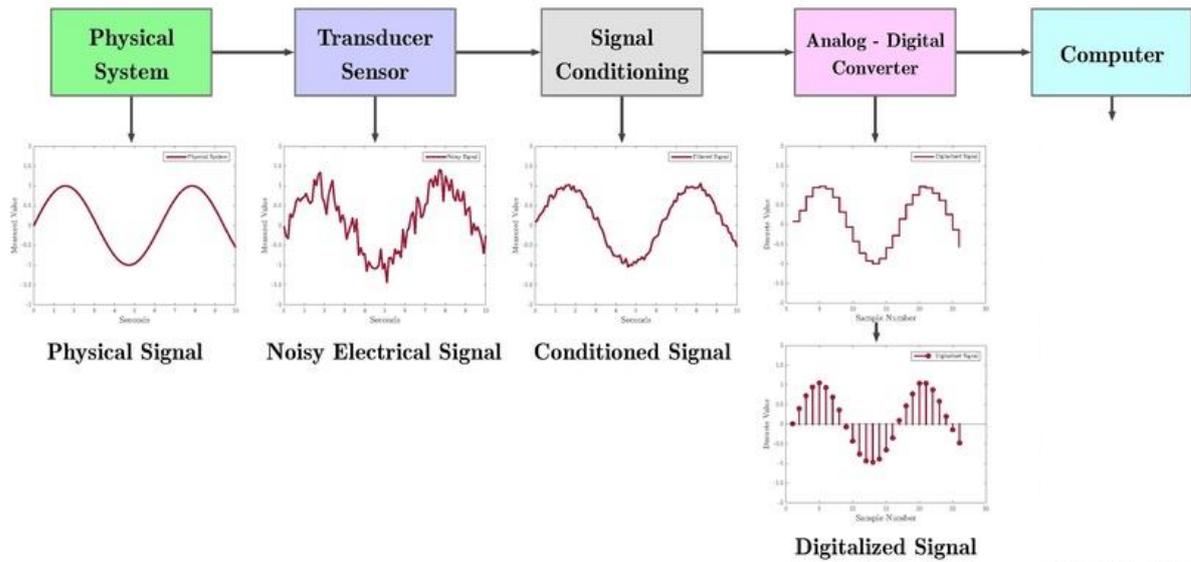
In questo corso ci occuperemo per lo più della parte di conversione dei dati e dell'analisi dei dati.

Un altro schema è il seguente¹⁰:

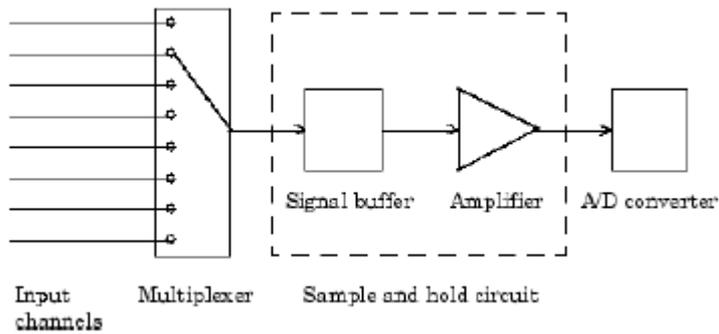
⁹ Dal manuale DAQ User Guide di Matlab.

¹⁰ Da "Data acquisition" in Wikipedia English edition.

Digital Data Acquisition System



Spesso un convertitore analogico digitale viene utilizzato per più canali di acquisizione:



Teorema del campionamento e aliasing

Il problema è questo: se campioniamo un segnale discreto, ne cogliamo tutte le caratteristiche? In altre parole, possiamo, a partire dai campioni, ricostruire il segnale originale?

Il teorema del campionamento, formulato nel 1928 da Nyquist e provato nel 1949 da Shannon, è il seguente:

Se un segnale reale ha spettro di potenza nullo per frequenze superiori a una data ν_0 , esso può essere completamente ricostruito se è campionato ad una frequenza

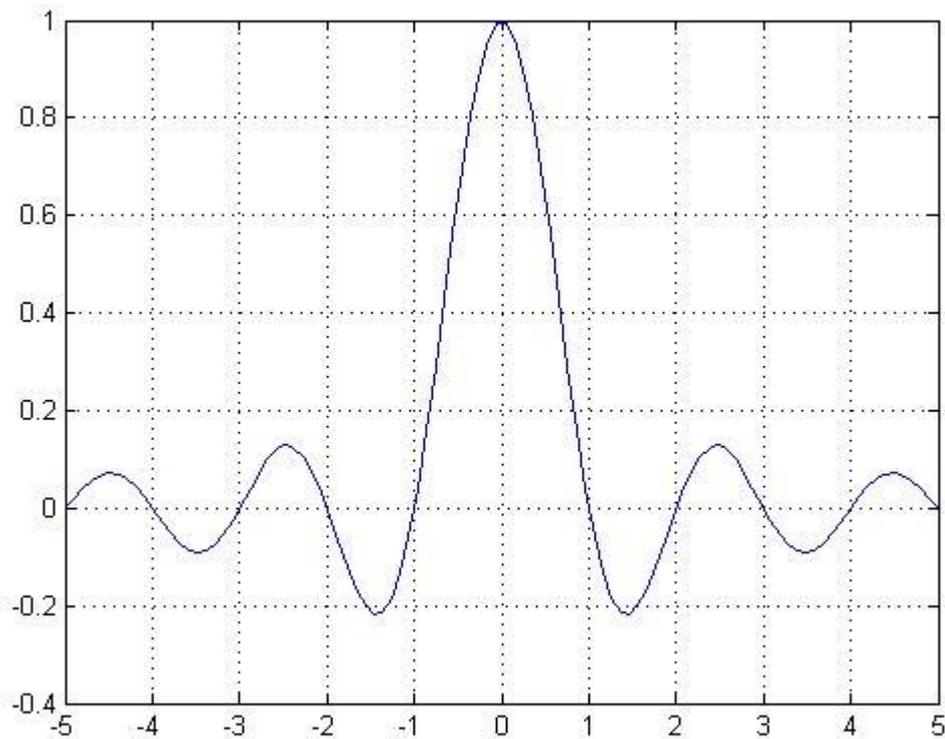
$$(5.1) \quad \nu_s \geq 2 \cdot \nu_0$$

Data una frequenza di campionamento $\nu_s = \frac{1}{\Delta t_s}$, dove Δt_s è il tempo di campionamento, la frequenza $\nu_N = \frac{\nu_s}{2}$ è detta **frequenza di Nyquist**.

Dati i campioni $\{x_i\}$, nell'ipotesi del teorema, possiamo ricavare il segnale originale come

$$(5.2) \quad x(t) = \sum_{i=-\infty}^{\infty} x_i \cdot \frac{\sin\left(\pi\left(\frac{2t}{\Delta t_s} - i\right)\right)}{\pi\left(\frac{2t}{\Delta t_s} - i\right)}$$

la funzione $\text{sinc}(\alpha) = \frac{\sin(\pi \cdot \alpha)}{\pi \cdot \alpha}$ che compare nella (5.2), detta funzione interpolante, ha il seguente andamento



ed è la trasformata di Fourier della funzione

$$(5.3) \quad f(t) = \begin{cases} \frac{1}{2\pi} & \text{per } -\pi \leq t < \pi \\ 0 & \text{altrove} \end{cases}$$

Osservando la funzione d'interpolazione, notiamo che correttamente essa si annulla nei punti in cui ci sono altri campioni (infatti a quegli istanti la $x(t)$ è completamente definita dal campione), ma può stupire che in certi punti sia addirittura negativa: come mai? La ragione è dovuta alle correlazioni negative introdotte dal taglio in frequenza.

Notiamo che la (5.2) non è altro che l'uscita di un filtro passa-basso "perfetto", con funzione di trasferimento

$$(5.4) \quad F(j\omega) = \begin{cases} 1 & \text{per } -2\pi \cdot \nu_N \leq \omega < 2\pi \cdot \nu_N \\ 0 & \text{altrove} \end{cases}$$

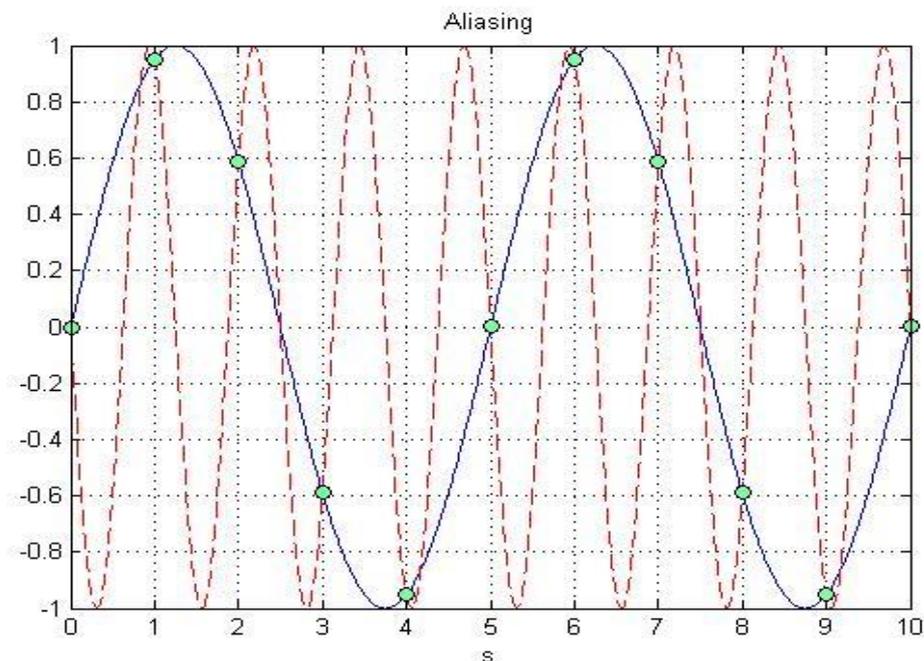
al cui ingresso sia stato posto il segnale

$$(5.5) \quad x(t) = \sum_{i=-\infty}^{\infty} x_i \cdot \delta(t - i \cdot \Delta t_s)$$

Ma che succede se si campiona con un tempo di campionamento che non rispetta la condizione posta dall'equazione (5.1) del teorema del campionamento ?

Ciò che succede è che le frequenze più elevate di $\nu_N = \frac{\nu_s}{2}$, frequenza di Nyquist, si “mascherano” da frequenze inferiori a ν_N . Questo fenomeno viene detto con termine inglese **aliasing**.

In figura sono mostrate due sinusoidi, una a frequenza $\nu = \frac{1}{5} \text{ Hz}$ (la frequenza di campionamento è 1 Hz) e una a $\nu' = \nu_s - \nu = \frac{4}{5} \text{ Hz}$.



Si vede che le due sinusoidi condividono esattamente gli stessi campioni, anche se la sinusoida a più bassa frequenza ha 5 campioni a periodo, quella a più alta 1.25 campioni (meno dei 2 prescritti dal teorema del campionamento).

In genere una frequenza

$$(5.6) \quad \nu = N \cdot \nu_s + \nu_1 \quad \text{essendo} \quad \nu_1 < \nu_s$$

si “maschera” da ν_1 se $\nu_1 < \frac{\nu_s}{2}$, altrimenti si maschera da $\nu_s - \nu_1$, o, più correttamente, da $-\nu_1$.

Il fenomeno dell’aliasing è ben noto nella vita pratica. Per esempio è dovuto all’aliasing il fatto che a volte nei film le ruote delle carrozze sembrano girare al contrario (un film è infatti una successione di immagini campionate, in genere a 24, 25 o 30 Hz, se i raggi delle carrozze hanno tra di loro un angolo α e tra un fotogramma e l’altro la ruota gira di un po’ più di $\frac{\alpha}{2}$, questa appare girare al contrario. Accelerando, apparirà girare nel verso giusto, ma più lentamente, quindi di nuovo nell’altro verso e così via.

Ci sono usi pratici importanti dell’aliasing, per esempio lo stroboscopio.

Il fenomeno dell’aliasing è in genere dannoso nell’analisi dei segnali perché aumenta il livello del rumore. Va quindi eliminato o per lo meno limitato, e ciò si fa con opportuni filtri passa basso, **analogici**, prima del campionamento, in modo da porsi nelle condizioni del teorema del campionamento (nessuna componente a frequenze superiori alla frequenza di Nyquist).

Talvolta, data la molto maggiore facilità di realizzazione dei filtri digitali, se si vuole un campionamento a una frequenza ν_s , si usa un “rozzo” filtro anti-aliasing analogico, una frequenza di campionamento molto più elevata di quella voluta, quindi un filtro anti-aliasing digitale “perfetto” che limita il segnale nella banda di Nyquist voluta. Si procede quindi a un sottocampionamento alla frequenza voluta. Questa procedura offre anche il vantaggio di ridurre drasticamente il rumore di quantizzazione.

Infine vogliamo ricordare che talvolta l’aliasing nell’analisi dei segnali può essere usato utilmente, per esempio per abbassare la frequenza di campionamento (vedi il prossimo paragrafo).

Generalizzazione

Se il segnale ha lo spettro nullo al di fuori di una certa banda B (che non necessariamente inizia da 0), possiamo teoricamente generalizzare il teorema del

campionamento, limitando la frequenza di campionamento a $2 \cdot B$. Ben più difficile è però la procedura di ricostruzione.

Data una certa frequenza di campionamento, abbiamo visto che l'asse delle frequenze si suddivide in bande "aliasate": se la banda del segnale entra tutta in una di queste bande aliasate, allora la ricostruzione è semplificata. Ciò però può significare dover scegliere una frequenza di campionamento abbastanza maggiore di $2 \cdot B$.

Il caso dei segnali complessi

Se il segnale è complesso, c'è differenza tra le frequenze positive e negative, quindi la banda del segnale da considerare è doppia. Tuttavia il numero di campioni necessari è la metà, perché ciascun campione è "doppio" (parte reale e parte immaginaria). Detto in altre parole, abbiamo il doppio di campioni reali, ma descriviamo una banda doppia: la simmetria rispetto alla frequenza di Nyquist non c'è.

Il rumore di quantizzazione

Quando si digitalizza un segnale continuo, oltre ai problemi relativi al campionamento, si devono considerare i problemi relativi alla quantizzazione. Il segnale campionato infatti viene trasformato in un numero intero (e memorizzato come tale) tramite l'uso di un convertitore analogico-digitale (ADC). Questo numero è il numero che, moltiplicato per una quantità Δx chiamata **quanto di conversione**, meglio approssima il valore del campione (che è un numero reale). Se il campione vale x , ricaviamo un valore n tale che

$$(5.7) \quad -\frac{\Delta x}{2} < x - n \cdot \Delta x \leq \frac{\Delta x}{2}$$

Al valore vero x del campione assegneremo il valore "quantizzato" $x = n \cdot \Delta x = x + \varepsilon$.

Possiamo schematizzare questo processo come l'aggiunta di un rumore ε , detto **rumore di quantizzazione**.

Se il segnale $\{x_i\}$ è schematizzabile con un rumore bianco (cioè i campioni sono indipendenti) i campioni $\{\varepsilon_i\}$ del rumore di quantizzazione sono indipendenti e distribuiti uniformemente tra $-\frac{\Delta x}{2}$ e $\frac{\Delta x}{2}$.

La varianza è

$$(5.8) \quad \sigma_\varepsilon^2 = \frac{(\Delta x)^2}{12}$$

e lo spettro di potenza è piatto e, per il teorema di Parseval,

$$(5.9) \quad S(\omega) = \frac{(\Delta x)^2}{12} \cdot \Delta t$$

costante, indipendente da ω .

Si noti che la densità spettrale decresce con l'aumentare della frequenza di campionamento. Ciò può essere utilizzato per ridurre l'errore di quantizzazione: vediamo come.

La “classe” di un ADC è data dal numero di bit N usati per la conversione, in genere 8 o 12 o 16. Quanto maggiore è il numero di bit, tanto migliore è il convertitore: essendo M il massimo range coperto dal convertitore, il quanto di conversione è

$$(5.10) \quad \Delta x = \frac{M}{2^N}$$

L’idea è quella di campionare i dati analogici a una frequenza v_s' molto più elevata di quella richiesta v_s , applicare un buon filtro anti-aliasing (in software) e sotto-campionare a v_s : i campioni ottenuti ora sono numeri “reali” (non più “interi”), perché l’uscita del filtro è composta da numeri floating point (a 32 bit, che equivale a circa 24 bit di mantissa, o più se si lavora in doppia precisione) e avranno un errore di

quantizzazione equivalente con spettro $\frac{v_s'}{v_s}$ volte più basso e quindi è come se avessimo

$$(5.11) \quad \frac{1}{2} \log_2 \left(\frac{v_s'}{v_s} \right)$$

bit di conversione in più.

Signal conditioning

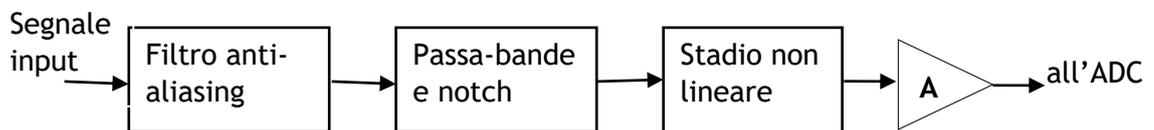
Abbiamo visto quali sono i principali problemi che sorgono nel campionare un segnale: dai dati campionati si devono ricavare il maggior numero di informazioni, non dobbiamo distorcere il segnale e dobbiamo minimizzare il rumore.

Prima di campionare un segnale, dobbiamo definire quale è il range di frequenze di interesse, definendo soprattutto qual è la massima frequenza di interesse v_{\max} e quindi scegliendo la frequenza di campionamento in modo tale che sia $v_s = 2 \cdot v_N > 2 \cdot v_{\max}$ (v_N è la frequenza di Nyquist). Realizzeremo quindi un filtro passa basso (detto filtro antialiasing) che riduca molto il rumore a frequenze superiori alla frequenza di Nyquist: questo rumore infatti si somma a quello nella banda.

È utile anche definire se ci sono altre bande di frequenza in cui il segnale di nostro interesse è trascurabile o il rumore è particolarmente alto. Un caso tipico è la frequenza a 50 Hz (60 negli Stati Uniti e altrove) e sue armoniche dove spesso si trovano forti disturbi dovuti alla rete elettrica: talvolta ci si mette un filtro “notch” (che cancella una banda molto stretta).

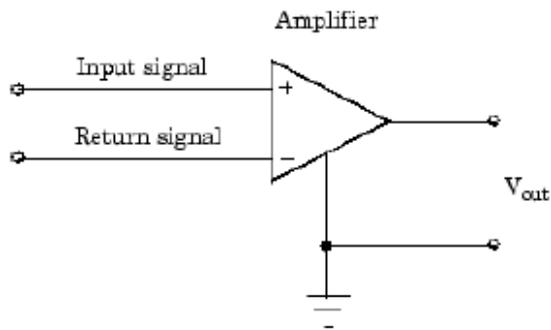
Talvolta inoltre si inserisce uno stadio non lineare che permette di “comprimere” la dinamica di ampiezza.

Infine si pone uno stadio di amplificazione (o riduzione) che adatta la dinamica del segnale al range di conversione dell'ADC, in modo da descrivere meglio il segnale, ovvero ridurre il rumore di quantizzazione (relativo).



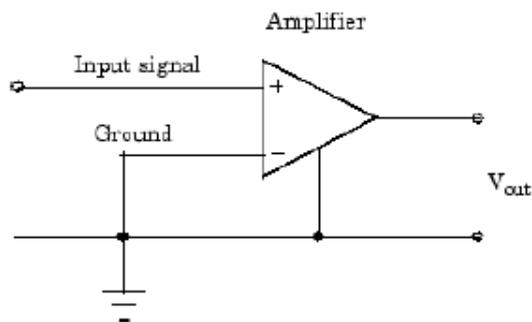
Se il segnale è particolarmente debole occorre preamplificarlo (prima degli altri stadi) facendo attenzione a non aggiungere ulteriore rumore con i cavi di connessione. Il preamplificatore è spesso un amplificatore differenziale capace di cancellare gli eventuali disturbi raccolti dai cavi di connessione.

Consigli della National Instrument:



National Instruments recommends that you use differential inputs under any of these conditions:

- The input signal is low level (less than 1 volt).
- The leads connecting the signal are greater than 10 feet.
- The input signal requires a separate ground-reference point or return signal.
- The signal leads travel through a noisy environment.



National Instruments suggests that you can use single-ended inputs under any of these conditions:

- The input signal is high level (greater than 1 volt).
- The leads connecting the signal are less than 10 feet.
- The input signal can share a common reference point with other signals.

You should use differential input connectors for any input signal that does not meet the preceding conditions. You can configure many National Instruments boards for two different types of single-ended connections:

6 - Sistemi discreti e modelli

Live Script

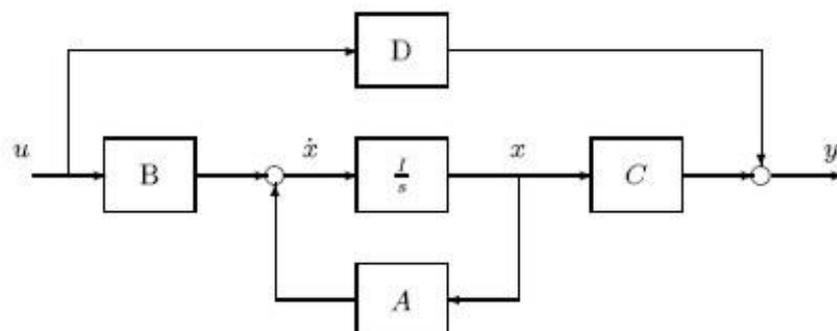
Generalità sui sistemi

Un sistema è un modello matematico di un dispositivo o di un processo fisico o di un algoritmo, che connette uno o più segnali di ingresso (eccitazione o “input”) a uno o più segnali di uscita (risposta o “output”). Un sistema quindi è sostanzialmente un elaboratore di segnali.



Mentre i processi fisici sono modellizzabili più naturalmente con sistemi continui ed i segnali fisici con segnali continui, il loro trattamento con i computer digitali rende fondamentale lo studio di sistemi e segnali discreti.

Più sistemi possono essere connessi insieme, e questo insieme può essere rappresentato da un nuovo sistema. Nella figura seguente è rappresentato un insieme di sistemi connessi tra di loro in vario modo.

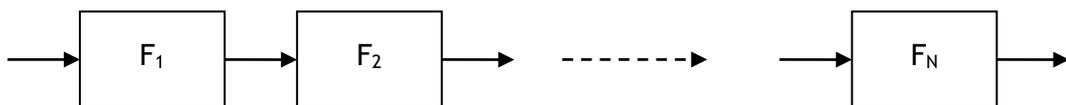


✍

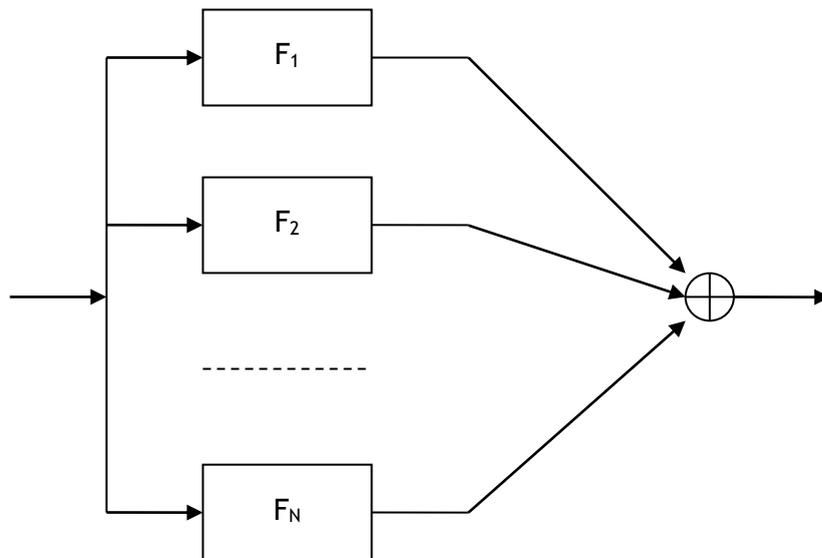
Si notino i simboli circolari che indicano i “nodi somma”, semplici sistemi con due o più ingressi e una sola uscita che eseguono la somma dei segnali di ingresso.

Modi particolarmente semplici, e spesso utili, di connettere dei sistemi sono i due seguenti:

- Connessione in serie (o in cascata)



- Connessione in parallelo



Attenzione ! Nella teoria dei sistemi il comportamento di un sistema non è modificato dalla presenza degli altri a cui è connesso. Come è noto, non è così per un normale sistema fisico (basti pensare ad un circuito RC passa-basso, la cui risposta dipende dall'impedenza d'ingresso del sistema che lo segue). Per sistemi elettrici, in cui le variabili di ingresso e di uscita sono tensioni elettriche, questa caratteristica equivale ad una impedenza d'ingresso infinita e un'impedenza di uscita nulla.

Come vedremo in questo corso, segnali e sistemi sono correlati tra di loro: i sistemi non ci serviranno solo per elaborare segnali, ma, in certi casi, per rappresentarli.

Classificazione dei sistemi

Oltre alla classificazione tra sistemi continui e discreti (cioè che elaborano segnali discreti o continui), esistono altre caratteristiche peculiari di un sistema:

- Statico o dinamico, detto anche senza memoria o con memoria, cioè se l'uscita ad un dato istante dipende solo dall'ingresso a quell'istante o no. Un sistema dinamico è caratterizzato dalla presenza di stati interni. In altri termini l'effetto della "memoria" del sistema viene schematizzato supponendo che l'uscita a un dato istante non dipende solo dall'ingresso a quell'istante, ma anche dal valore dello "stato". Quindi, mentre un sistema statico è completamente definito dall'equazione

$$(6.1) \quad \vec{y} = F(\vec{x})$$

dove \vec{x} e \vec{y} sono rispettivamente il vettore degli ingressi e quello delle uscite, ad un dato istante. Se il sistema è dinamico, invece, nel caso di sistemi continui, occorrono le due equazioni

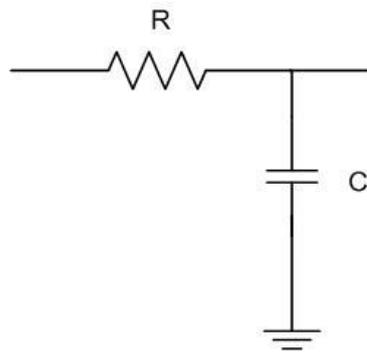
$$(6.2) \quad \begin{cases} \vec{y} = F(\vec{x}, \vec{s}) \\ \dot{\vec{s}} = G(\vec{x}, \vec{s}) \end{cases}$$

dove \vec{s} è il vettore di stato in quell'istante. La prima equazione viene detta equazione di uscita (output equation), la seconda equazione di stato (state equation). Nel caso di sistemi discreti,

$$(6.3) \quad \begin{cases} \vec{y}_i = F(\vec{x}_i, \vec{s}_i) \\ \vec{s}_i = G(\vec{x}_i, \vec{s}_{i-1}) \end{cases}$$

Un esempio di sistema dinamico è un circuito RC passa-basso, in cui l'ingresso e l'uscita siano tensioni: in tal caso la variabile di stato è la carica del condensatore¹¹.

¹¹ Il circuito in figura non ha le caratteristiche ideali di un sistema: infatti il suo "comportamento" varia a seconda di cosa mettiamo alla sua uscita ed inoltre esso "carica" il generatore che mettiamo all'ingresso. Se supponiamo che all'ingresso ci sia un generatore di



- Causale, se l'uscita non precede l'ingresso. Tutti i sistemi fisici sono causali, non così quelli simulati su computer.
- Lineare, se ad una qualsiasi combinazione lineare di differenti ingressi, corrisponde in uscita la stessa combinazione lineare delle relative uscite (principio di sovrapposizione).

Indichiamo con L l'operatore che descrive il sistema. Se il sistema è lineare, allora

$$(6.4) \quad L[a_1 \cdot x_1(t) + a_2 \cdot x_2(t)] = a_1 \cdot L[x_1(t)] + a_2 \cdot L[x_2(t)]$$

Nel caso di sistemi lineari, le equazioni (6.2) e (6.3) sono lineari.

- Tempo-invariante, se il comportamento non varia nel tempo. In tal caso le funzioni F e G sono indipendenti dal tempo.
- Stabile, se ad un ingresso di ampiezza limitata, corrisponde un'uscita di ampiezza limitata¹².

impedenza di uscita nulla e all'uscita un apparato di impedenza di ingresso infinita, il suo comportamento è assimilabile a un sistema ideale.

¹² Nella teoria dei sistemi di controllo vengono definiti vari tipi di stabilità che non interessano la presente breve trattazione.

Nel seguito ci occuperemo quasi esclusivamente di sistemi discreti lineari tempo invariati.

Sistemi discreti

Un sistema discreto (o a dati campionati) è un elaboratore di segnali discreti (o campionati). Cioè un sistema che accetta in ingresso una o più successioni di dati campionati e produce in uscita una o più successioni di dati campionati. Analogamente a quanto accade per i sistemi continui, che si possono definire in generale tramite l'equazione (6.2), essi possono definirsi tramite l'equazione (6.3), che riportiamo:

$$(6.5) \quad \begin{cases} \vec{y}_i = F(\vec{x}_i, \vec{s}_i) \\ \vec{s}_i = G(\vec{x}_i, \vec{s}_{i-1}) \end{cases}$$

dove \vec{x}_i , \vec{y}_i e \vec{s}_i sono rispettivamente i vettori d'ingresso, di uscita e di stato all'istante (campionato) i . Le dimensioni dei suddetti vettori sono in genere differenti; nel seguito ci occuperemo essenzialmente di sistemi con un ingresso ed una uscita, cioè con dimensioni di x e y uguali ad 1.

Se le funzioni F e G non dipendono dal tempo (cioè dall'indice i), il sistema è tempo-invariante. Nel seguito considereremo quasi esclusivamente sistemi tempo-invarianti (detti anche stazionari): trascureremo quindi di specificarlo.

Diciamo che un sistema è lineare se le funzioni F e G sono lineari, cioè sono equazioni matriciali. In tal caso

$$(6.6) \quad \begin{cases} \vec{y}_i = \mathbf{A} \cdot \vec{x}_i + \mathbf{B} \cdot \vec{s}_i \\ \vec{s}_i = \mathbf{C} \cdot \vec{x}_i + \mathbf{D} \cdot \vec{s}_{i-1} \end{cases}$$

con A , B , C e D delle matrici.

Ci sono altre rappresentazioni dei sistemi discreti lineari.

Definiamo la risposta impulsiva w_k del sistema come la risposta alla successione

$$(6.7) \quad \delta_i = \begin{cases} 0 & \text{per } i \neq 0 \\ 1 & \text{per } i = 0 \end{cases}$$

che è detta funzione impulsiva discreta, o impulso unitario o anche delta discreta ed ha nei sistemi discreti un ruolo analogo alla delta di Dirac per i sistemi continui.

Analogamente a quanto accade per i sistemi continui, per i sistemi lineari discreti (tempo-invarianti) la risposta a un ingresso generico x_i è data dalla convoluzione discreta della x_i per la risposta impulsiva w_k :

$$(6.8) \quad y_i = \sum_{k=-\infty}^{\infty} x_{i-k} \cdot w_k$$

Ovviamente nella (6.8), se la x_i è data dalla (6.7), allora $y_i = w_i$.

Se il sistema è causale, la (6.8) diventa

$$(6.9) \quad y_i = \sum_{k=0}^{\infty} x_{i-k} \cdot w_k$$

L'operazione di convoluzione discreta (6.8) viene spesso indicata in modo abbreviato, similmente alla convoluzione continua, con

$$(6.10) \quad y = x \otimes w$$

Se $w_k = 0$ per tutti i $k < 0$, il sistema è causale.

La (6.8) è un altro modo di rappresentare un sistema lineare, in genere più pratico di quello dell'equazione (6.6). In esso non compare lo stato, ma gli ingressi a tutti i tempi precedenti. Il vettore di stato può essere visto come la condizione in cui è stato pilotato il sistema dagli input precedenti.

Casi particolari

Ci sono poi casi particolari di sistemi lineari discreti, spesso molto utili:

- sistemi moving average (MA), cioè “a media mobile”, detti anche FIR (Finite Impulse Response), rappresentati dall’equazione

$$(6.11) \quad y_i = \sum_{k=0}^m b_k \cdot x_{i-k}$$

cioè l’uscita y_i è ottenuta dalla convoluzione delle successioni x (infinita) e b (finita). Come si vede immediatamente, questi sistemi sono caratterizzati da una risposta impulsiva composta da soli $m+1$ campioni. Hanno quindi una memoria finita: l’uscita è indipendente completamente da ciò che gli è stato dato in ingresso prima degli ultimi $m+1$ campioni.

- sistemi auto-regressivi (AR), detti anche IIR (Infinite Impulse Response), rappresentati dall’equazione

$$(6.12) \quad y_i = b_0 \cdot x_i - \sum_{k=1}^n a_k \cdot y_{i-k}$$

in cui l’uscita a un dato istante viene espressa in funzione dell’ingresso a quell’istante e da una combinazione delle uscite negli n istanti precedenti. Si dimostra che la risposta impulsiva di questo tipo di sistemi è di lunghezza infinita. Facciamo un esempio:

consideriamo il caso in cui $n=1$ e $a_1 = -w$. La (6.12) diventa

$$(6.13) \quad y_i = b_0 \cdot x_i + w \cdot y_{i-1}$$

e, ponendo come ingresso la successione (6.7), abbiamo

$$\begin{aligned}
 (6.14) \quad & y_i = 0 \quad \text{per } i < 0 \\
 & y_0 = b_0 \\
 & y_1 = b_0 \cdot w \\
 & y_2 = b_0 \cdot w^2 \\
 & y_3 = b_0 \cdot w^3 \\
 & \dots
 \end{aligned}$$

e quindi, in generale,

$$(6.15) \quad y_i = b_0 \cdot w^i$$

per qualsiasi $i > 0$.

- sistemi ARMA (auto-regressive moving average), rappresentati dall'equazione

$$(6.16) \quad y_i = \sum_{k=0}^m b_k \cdot x_{i-k} - \sum_{k=1}^n a_k \cdot y_{i-k}$$

che comprende una combinazione lineare degli $m+1$ ingressi precedenti e le n uscite precedenti. Generalizza i due casi precedenti.

In Matlab la funzione filter realizza un generico filtro ARMA.

Semplice applicazione di sistemi MA e AR

Supponiamo di avere una successione x_i di misure di una grandezza fisica s costante, "disturbate" da un "rumore" (cioè da errori casuali),

$$(6.17) \quad x_i = s + n_i$$

Supponiamo che gli errori n_i siano tra loro indipendenti e di voler stimare il valor medio di s . Come è noto, un buon metodo per fare questa stima è prendere la media di un

certo numero di valori delle misure x_i . Supponiamo che gli errori (o il “rumore”) n_i abbiano valor medio nullo e deviazione standard σ ; possiamo definire σ “incertezza”¹³ su s da una singola misura. Se facciamo la media su N misure, diminuiamo l’incertezza su s di \sqrt{N} volte.

Supponiamo ora che la grandezza s vari lentamente nel tempo. Come possiamo stimarla? Un modo potrebbe essere prendere la media di un numero N di misure successive abbastanza piccolo che, durante le N , s sia variato in modo trascurabile¹⁴. Ciò significa che possiamo fare ad ogni istante i una stima μ_i del valor medio di s . Abbiamo

$$(6.18) \quad \mu_i = \frac{1}{N} \sum_{k=0}^{N-1} x_{i-k}$$

Ciò è realizzato tramite un sistema MA con coefficienti tutti eguali $b_k = \frac{1}{N}$, per $0 \leq k \leq N-1$.

Un problema di questo modo di procedere è che appare piuttosto innaturale considerare tutti e soli gli ultimi N campioni, tutti con lo stesso peso. Si possono fare scelte dei pesi che decrescono con k , in modo che i valori più lontani, e che quindi possono avere un s più diverso rispetto al valore all’istante i , pesino di meno.

Una scelta particolarmente “ragionevole” è “comoda” è usare un sistema AR del tipo di eq. (6.13), con $w = e^{-\frac{1}{N}}$. Ciò equivale a fare una media mobile con infiniti termini (in linea di principio, in pratica si considerano tutti i termini precedenti disponibili), con pesi decrescenti in modo esponenziale con la distanza. Si noti che questa soluzione è anche molto “leggera” computazionalmente, perché ad ogni passo richiede solo una moltiplicazione e una somma¹⁵.

¹³ Potremo dare differenti definizioni dell’incertezza, per esempio 2σ , o 3σ .

¹⁴ Vedremo in seguito come le varie espressioni qualitative presenti in questo paragrafo possano essere rese quantitative.

¹⁵ Anche un filtro MA con tutti i pesi uguali può realizzarsi con un basso costo computazionale, ma con un maggiore “costo” di memoria.

Abbiamo qui considerato solo sistemi “causali” che permettono di dare la stima μ_i non appena è disponibile il campione x_i . In molti casi non c'è questa restrizione, e quindi per μ_i possono utilizzarsi anche i campioni con indice maggiore di i .

In particolare il filtro MA visto precedentemente diventa, ovviamente con un diverso N ,

$$(6.19) \quad \mu_i = \frac{1}{2N+1} \sum_{k=-N}^N x_{i-k}$$

Per quanto riguarda il filtro AR, possiamo applicarlo due volte, una volta (“in avanti”) normalmente e l'altra (“all'indietro”) invertendo l'ordine dei campioni; in tal modo il filtraggio avviene considerando per ciascun campione tutti i campioni, sia precedenti che seguenti.

Equazioni alle differenze

L'equazione (6.16) può essere riscritta come

$$(6.20) \quad \sum_{k=0}^n a_k \cdot y_{i-k} = \sum_{k=0}^m b_k \cdot x_{i-k}$$

dove $a_0 = 1$. Sottolineiamo che questa equazione descrive un sistema causale¹⁶.

Un'equazione come la (6.20) è detta equazione lineare alle differenze.

Si noti la somiglianza con il caso della generica equazione differenziale lineare:

$$(6.21) \quad \sum_{k=0}^n A_k \frac{d^k y(t)}{dt^k} = \sum_{l=0}^m B_l \frac{d^l x(t)}{dt^l}$$

in questa l'operatore differenziale sostituisce l'operatore ritardo presente in quella. E infatti nei sistemi discreti l'operatore "avanzamento" ha una funzione analoga all'operatore differenziale dei sistemi continui: le equazioni alle differenze finite, come la (6.20), prendono il posto delle equazioni differenziali.

Si noti che i due termini della (6.20) possono essere visti come due convoluzioni, della sequenza finita $\{a_i\}$ e della sequenza infinita $\{y_i\}$ e della sequenza finita $\{b_i\}$ e della sequenza infinita $\{x_i\}$. Definendo le trasformate z

$$(6.22) \quad \begin{aligned} X(z) &= \sum_{i=-\infty}^{\infty} x_i \cdot z^{-i} \\ Y(z) &= \sum_{i=-\infty}^{\infty} y_i \cdot z^{-i} \\ A(z) &= \sum_{i=0}^n a_i \cdot z^{-i} \\ B(z) &= \sum_{i=0}^m b_i \cdot z^{-i} \end{aligned}$$

¹⁶ Questa trattazione è valida per sistemi causali, ma può facilmente adattarsi a sistemi non causali, con la limitazione che a) l'uscita dipenda da ingressi posteriori, ma non da uscite posteriori, b) ci sia un limite al massimo "futuro". Il caso di sistemi anti-causali è ovviamente banale.

possiamo utilizzare la proprietà della trasformata z della convoluzione ed otteniamo l'equazione

$$(6.23) \quad A(z)Y(z) = B(z)X(z)$$

Funzione di trasferimento discreta

Riscriviamo l'equazione (6.23) come

$$(6.24) \quad Y(z) = \frac{B(z)}{A(z)} \cdot X(z)$$

e quindi l'uscita y del sistema (o meglio la sua trasformata z) può ottenersi dall'ingresso x , z -trasformato e moltiplicato per la funzione

$$(6.25) \quad F(z) = \frac{B(z)}{A(z)}$$

o anche, esplicitando,

$$(6.26) \quad F(z) = \frac{\sum_{i=0}^m b_i \cdot z^{-i}}{1 + \sum_{i=1}^n a_i \cdot z^{-i}}$$

La funzione $F(z)$ è detta funzione di trasferimento discreta e opera analogamente alla funzione di trasferimento definita per i sistemi continui.

Possiamo sviluppare la (6.26) in serie di potenze di z , ottenendo

$$(6.27) \quad F(z) = \sum_{i=0}^{\infty} f_i \cdot z^{-i}$$

(se $n=0$, la serie è finita), dove le $\{f_i\}$ sono la risposta impulsiva.

È chiaro che la (6.24) risolve in modo formalmente semplice il problema del calcolo della risposta forzata dell'equazione (6.20) (ma in pratica potrebbe risultare piuttosto complesso, dato il problema del calcolo della trasformata dell'ingresso e soprattutto dell'anti-trasformata dell'uscita).

La funzione di trasferimento è il modo più usato di rappresentare un sistema lineare discreto tempo-invariante.

Una delle proprietà più importanti della funzione di trasferimento z (simile a quella delle funzioni di trasferimento con la trasformata di Laplace) è che la funzione di trasferimento di un sistema composto da due sistemi in cascata F_1 e F_2 , ha come funzione di trasferimento il prodotto delle due:

$$(6.28) \quad F = F_1 \cdot F_2$$

Possiamo anche definire il sistema inverso di uno dato, tale che se messo in cascata con esso il sistema risultante ha funzione di trasferimento 1, semplicemente come

$$(6.29) \quad F_I(z) = \frac{1}{F(z)}$$

tuttavia, perché sia “utilizzabile”, dobbiamo verificare che $F_I(z)$ sia stabile.

Poiché la trasformata z della delta discreta è 1, la $F(z)$ è anche la trasformata z della risposta impulsiva del sistema.

Ricordando che sul cerchio unitario $z = e^{j\Omega}$, possiamo facilmente calcolare la risposta in frequenza del sistema (la risposta a una sinusoide campionata) come

$$(6.30) \quad \mathcal{F}(\Omega) = F(e^{j\Omega}) = \frac{\sum_{k=0}^m b_k \cdot e^{-j\Omega \cdot k}}{1 + \sum_{k=1}^n a_k \cdot e^{-j\Omega \cdot k}}$$

La pulsazione Ω è definita per $0 \leq \Omega \leq 2\pi$ e, se i coefficienti a e b sono reali,

$$(6.31) \quad F(e^{j\Omega}) = F^*(e^{-j\Omega})$$

e quindi il valore massimo “effettivo” di Ω è π . I valori di Ω tra π e 2π possono vedersi come frequenze angolari negative (ricordiamo che l’asse Ω è in effetti una circonferenza) e per sistemi a coefficienti reali la simmetria (6.31) è una simmetria tra frequenze positive e negative. Per i sistemi a coefficienti non reali, tale simmetria non sussiste.

Si noti che, nel caso dei coefficienti reali, per $\Omega = 0$ e $\Omega = \pi$ \mathcal{F} è reale, quindi lo sfasamento è nullo.

Usando la pulsazione (o frequenza angolare) “fisica” ω ,

$$(6.32) \quad \mathcal{F}(\omega \cdot \Delta t) = F(e^{j\omega \Delta t})$$

$\mathcal{F}(\Omega)$ è una funzione complessa della pulsazione normalizzata Ω : per ciascun valore di Ω abbiamo un numero complesso il cui modulo indica il guadagno del sistema e la fase lo sfasamento. Possiamo quindi rappresentare la $\mathcal{F}(\omega \cdot \Delta t)$ con i classici diagrammi di Bode.

Analogamente a come si fa per il continuo, la funzione di trasferimento è completamente descritta avendo dato

- gli zeri del polinomio al numeratore (chiamati semplicemente “zeri” z_k)
- gli zeri del polinomio al denominatore (chiamati “poli” p_k)
- il valore di b_0

Differenze e derivate

Data una successione $\{x_i\}$, chiamiamo differenza prima di $\{x_i\}$ la successione

$$(6.33) \quad y_i = x_i - x_{i-1}$$

La (6.33) rappresenta un sistema MA del primo ordine con funzione di trasferimento

$$(6.34) \quad F(z) = 1 - z^{-1}$$

Possiamo costruire la differenza seconda di $\{x_i\}$, facendo la differenza prima di $\{y_i\}$; il "sistema" che fa la differenza seconda equivale a due sistemi che fanno la differenza prima posti in serie, quindi la sua funzione di trasferimento è il quadrato della (6.34)

$$(6.35) \quad F(z) = (1 - z^{-1})^2 = 1 - 2z^{-1} + z^{-2}$$

Con le potenze successive della (6.34) si possono costruire le differenze di ordine superiore.

Si noti l'analogia della differenza prima (nel discreto) con la derivata prima (nel continuo).

Si può invertire l'operazione di differenza prima, con un sistema con funzione di trasferimento inversa

$$(6.36) \quad F(z) = \frac{1}{1 - z^{-1}}$$

che equivale al sistema AR del primo ordine

$$(6.37) \quad y_i = x_i + y_{i-1}$$

che esegue un'operazione, chiamata somma cumulativa, analoga a quella dell'integrale nel continuo

$$(6.38) \quad y(t) = \int_{-\infty}^t x(\vartheta) \cdot d\vartheta$$

Ma quanto l'operazione di differenza è vicina a quella di derivazione ? Questa questione potrà essere discussa quantitativamente alla luce di concetti che svilupperemo in seguito (lo spettro di potenza o di energia di un segnale), ma è chiaro che la differenza sarà una tanto migliore "stima" della derivata, quanto più lenta è la variazione del segnale (a parte un coefficiente dovuto al tempo di campionamento, che noi normalmente, lavorando direttamente nel discreto, supponiamo pari all'unità di tempo).

Risposta di un sistema discreto

La soluzione generale di un'equazione lineare alle differenze è data dalla somma dell'*evoluzione libera*, cioè la soluzione ottenuta con ingresso nullo, e dalla *risposta forzata*, cioè dalla soluzione ottenuta con le condizioni iniziali (gli n valori precedenti il primo valore di ingresso non nullo) tutte poste a 0.

Il caso più semplice di risposta forzata è la risposta impulsiva, che è la risposta forzata in cui la "forzante", cioè l'ingresso, è un segnale a delta unitario. A partire dalla risposta impulsiva si può ricavare la risposta forzata e a quest'ultima può ricondursi l'evoluzione libera.

Risposta impulsiva

La risposta impulsiva di un sistema descritto dalla (6.20) può essere ottenuta semplicemente in modo ricorsivo, ponendo $x_0 = 1$ e tutti i valori successivi a 0.

Un altro modo per ricavarla è eseguire la divisione tra i polinomi (nella variabile z^{-1}) a numeratore e a denominatore, ottenendo (in genere) una equivalenza tra la funzione di trasferimento e un polinomiale infinito, in cui il coefficiente della potenza i -esima è l' i -esimo valore della risposta impulsiva.

Per ottenere una soluzione in forma analitica, osserviamo che il sistema associato alla (6.20), descritto dalla funzione di trasferimento (6.26), nel caso in cui tutti gli n poli siano differenti, può essere visto come il parallelo di n sistemi, ciascuno dei quali ha un solo polo.

Quindi il k -esimo polo $w_{(k)}$ darà una componente additiva al risultato proporzionale a

$$(6.39) \quad \eta_i^{(k)} = w_{(k)}^i$$

avremo quindi per la risposta impulsiva complessiva

$$(6.40) \quad \eta_i = \sum_{k=1}^n K_k \cdot \eta_i^{(k)}$$

dove le costanti K_i possono essere calcolate dal sistema formato dalle prime n

$$(6.41) \quad \sum_{k=1}^n K_k \cdot \eta_i^{(k)} = \eta_i$$

dove i η_i sono i η_i calcolati in modo ricorsivo (e ovviamente $\eta_i = \eta_i$).

Risposta forzata

Se conosciamo la risposta impulsiva η_i , possiamo ottenere la risposta forzata, ad un ingresso $\{x_i\}$, come (vedi equazione (6.8))

$$(6.42) \quad y_i = \sum_{k=0}^{\infty} x_{i-k} \cdot \eta_k$$

Evoluzione libera

L'evoluzione libera si ha quando l'ingresso è nullo, ma non sono nulli tutti i valori di y_{i-k} per $1 \leq k \leq n$. Per l'evoluzione libera, possiamo considerare l'omogenea associata alla (6.20)

$$(6.43) \quad \sum_{k=0}^n a_k \cdot y_{i-k} = 0$$

o, ricordando $a_0 = 1$,

$$(6.44) \quad y_i = -\sum_{k=1}^n a_k \cdot y_{i-k}$$

A partire dalle condizioni iniziali

$$(6.45) \quad y_{-1}, y_{-2}, \dots, y_{-n}$$

possiamo calcolare ricorsivamente l'evoluzione libera come

$$(6.46) \quad \begin{aligned} y_0 &= -a_1 y_{-1} - a_2 y_{-2} - \dots - a_n y_{-n} \\ y_1 &= -a_1 y_0 - a_2 y_{-1} - \dots - a_n y_{-n+1} \\ y_2 &= -a_1 y_1 - a_2 y_0 - \dots - a_n y_{-n+2} \\ &\dots \end{aligned}$$

Possiamo tuttavia ricavare per essa un'espressione analitica. Per far ciò notiamo che l'evoluzione libera, nel caso delle condizioni (6.45), è equivalente alla risposta forzata della parte AR, in cui l'ingresso è dato da

$$(6.47) \quad \begin{aligned} x_0 &= -a_1 y_{-1} - a_2 y_{-2} - \dots - a_n y_{-n} \\ x_1 &= -a_2 y_{-1} - a_3 y_{-2} - \dots - a_n y_{-n+1} \\ x_2 &= -a_3 y_{-1} - a_2 y_0 - \dots - a_n y_{-n+2} \\ &\dots \\ x_{n-2} &= -a_{n-1} y_{-1} - a_n y_{-2} \\ x_{n-1} &= -a_n y_{-1} \end{aligned}$$

quindi basta calcolare questa risposta forzata.

Notiamo che l'evoluzione libera è indipendente dal numeratore della funzione di trasferimento (parte MA dell'equazione alle differenze).

Stabilità

Condizione necessaria e sufficiente per la stabilità è che il modulo dei poli della funzione di trasferimento sia minore di 1: i poli siano cioè interni al cerchio unitario.

Questa è la stabilità analitica. In pratica può accadere che un sistema analiticamente stabile non lo sia nella sua implementazione su computer. Ciò può accadere, per sistemi complessi con poli molto vicini all'unità, a causa degli errori di arrotondamento.

Notiamo che i sistemi MA (o FIR), non avendo poli, sono intrinsecamente stabili.

Sistemi semplici

Analizziamo ora in dettaglio alcuni semplici sistemi MA e AR, cioè sistemi in cui nell'equazione (6.20), che per comodità riscriviamo,

$$(6.48) \quad \sum_{k=0}^n a_k \cdot y_{i-k} = \sum_{k=0}^m b_k \cdot x_{i-k}$$

n (detto ordine AR) e m (detto ordine MA) hanno valori piccoli.

Spesso sistemi più complessi possono essere ricondotti a più di questi sistemi in cascata.

Sistema di ordine 0

Il sistema più semplice è quello in cui nell'equazione (6.48), $n=m=0$ e abbiamo

$$(6.49) \quad y_i = b_0 \cdot x_i$$

Questo sistema non fa altro che amplificare di un fattore b_0 il segnale in ingresso. Questa amplificazione è la stessa per tutte le frequenze (la funzione di trasferimento è semplicemente b_0).

Attenzione ! Se b_0 è complesso e $b_0 = |b_0| \cdot e^{j\varphi}$, abbiamo un'amplificazione complessa che può essere vista come un'amplificazione reale $|b_0|$ più uno sfasamento (eguale per tutte le frequenze) di φ .

Sistema MA del primo ordine

Consideriamo il sistema in cui nell'equazione (6.48), $n=0$ e $m=1$; abbiamo

$$(6.50) \quad y_i = b_0 \cdot x_i + b_1 \cdot x_{i-1}$$

La funzione di trasferimento z è

$$(6.51) \quad F(z) = b_0 + b_1 \cdot z^{-1}$$

e la risposta in frequenza è

$$(6.52) \quad \mathcal{F}(\Omega) = F(e^{j\Omega}) = b_0 + b_1 \cdot e^{-j\Omega}$$

La risposta impulsiva, finita, è composta da due termini:

$$(6.53) \quad \begin{aligned} x_0 &= b_0 \\ x_1 &= b_1 \end{aligned}$$

Consideriamo dei casi particolari.

- $b_0 = 0, b_1 = 1$, ovvero un sistema che esegue semplicemente un ritardo di un campione. Analizziamo la risposta in frequenza:

$$(6.54) \quad \mathcal{F}(\Omega) = e^{-j\Omega}$$

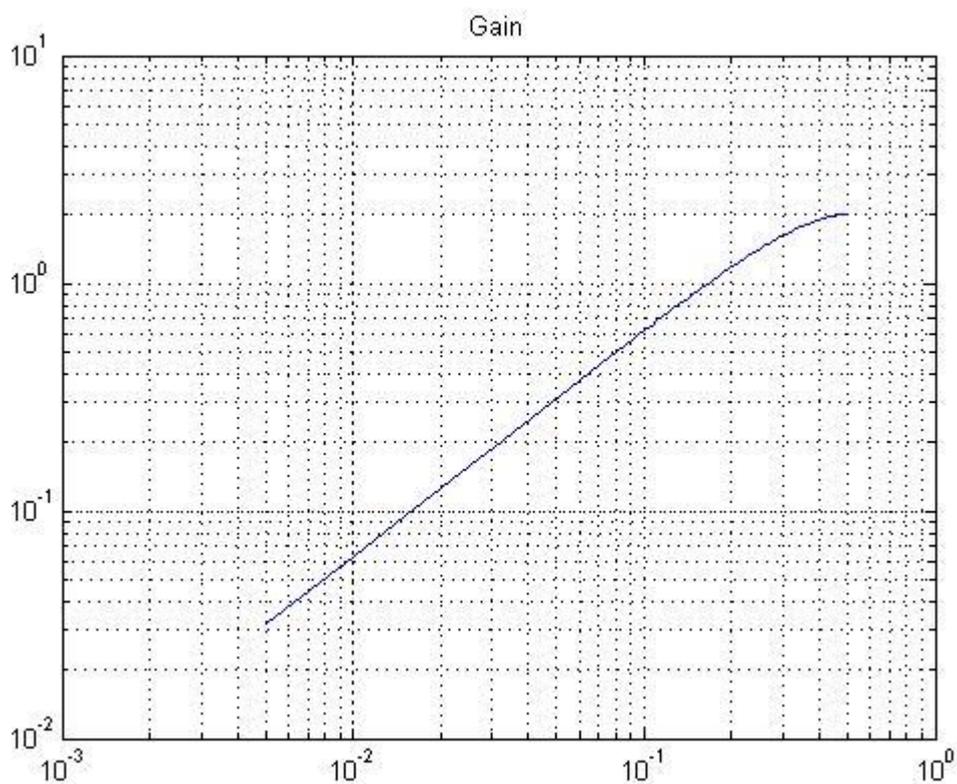
Vediamo che abbiamo un'amplificazione unitaria per tutte le frequenze e uno sfasamento proporzionale a Ω (si dice anche sfasamento lineare). È questa una caratteristica di qualsiasi ritardo.

Un sistema come questo, in cui il guadagno non varia con la frequenza, ma in cui varia lo sfasamento è detto sistema passa-tutto (all-pass system). Questo tipo di sistemi possono essere molto utili. Dato un sistema $F(z)$ possiamo sempre costruire un sistema passa-tutto con la stessa caratteristica di fase di $F(z)$, come

$$(6.55) \quad G(z) = \frac{F(z)}{|F(z)|}$$

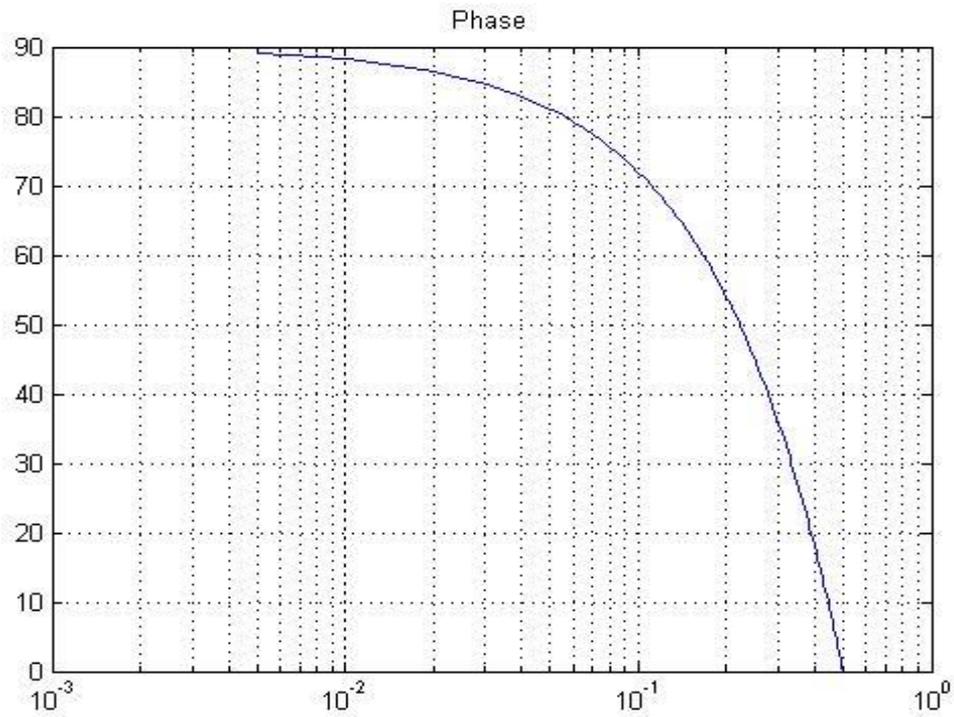
Tale sistema tuttavia in genere non sarà causale.

- $b_0 = -b_1 = 1$, ovvero un sistema che esegue la differenza prima, l'analogo discreto della derivata del continuo. Mostriamo i diagrammi di Bode¹⁷ per questo sistema:



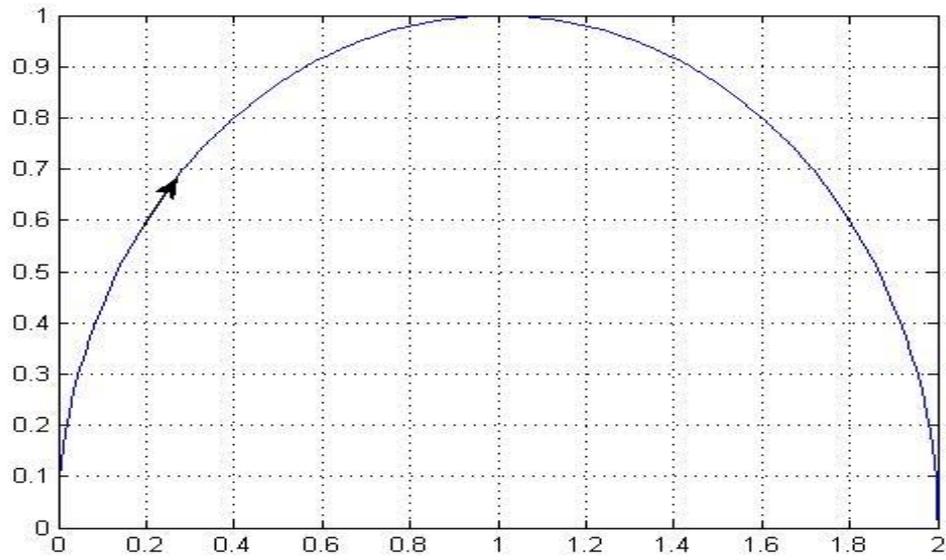
Si noti che la pendenza è di una decade di guadagno per una decade di frequenza.

¹⁷ Normalmente nei diagrammi di Bode il guadagno è espresso in decibel. Qui, per semplicità, lo lasciamo nelle sue unità naturali, in scala logaritmica. Ricordiamo che la misura in dB è $20 \cdot \log_{10} G$.



Si noti che a basse frequenze c'è uno sfasamento di 90 gradi positivo (in anticipo), mentre per alte frequenze lo sfasamento è nullo.

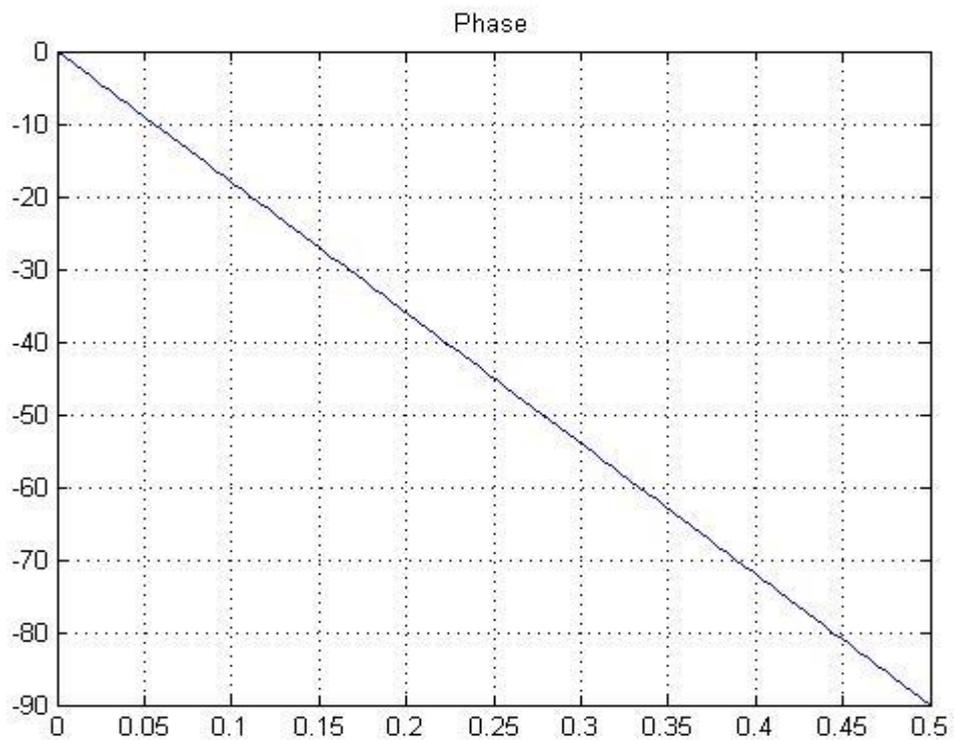
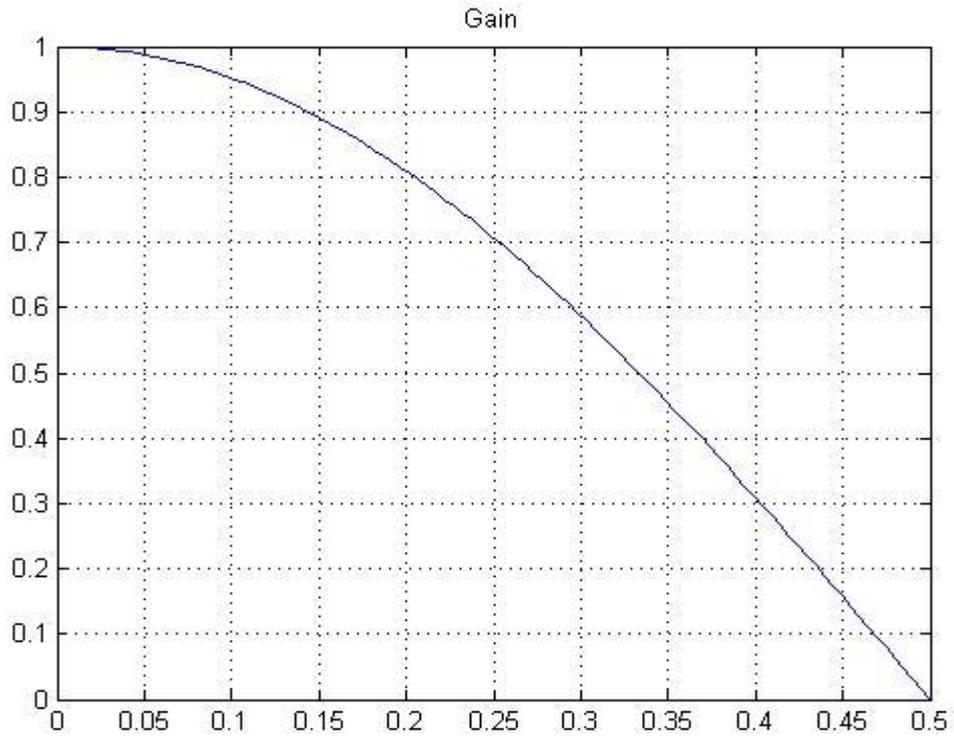
Possiamo anche graficare la funzione di trasferimento nel piano di Gauss (che ha per assi la parte reale e la parte immaginaria della variabile complessa), ottenendo il diagramma di Nyquist



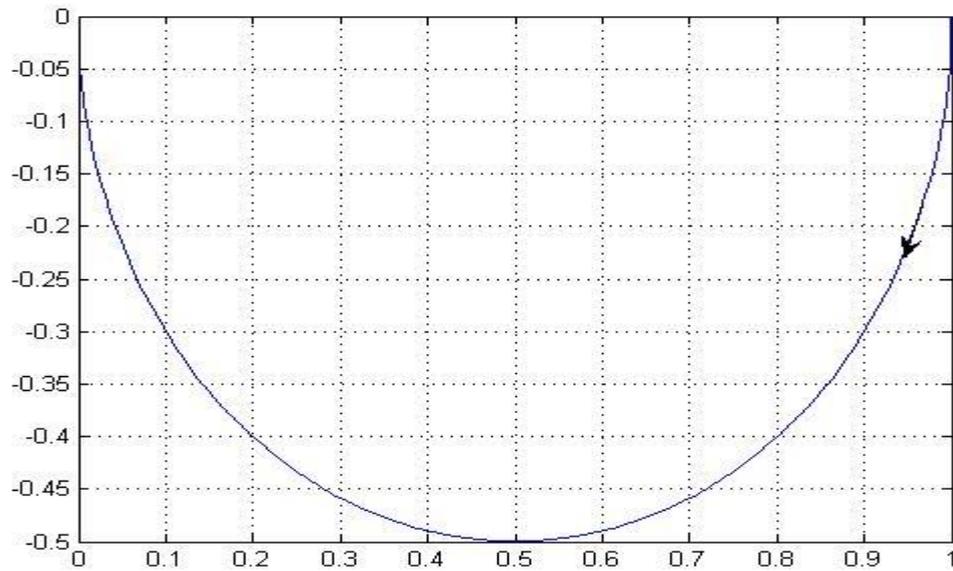
(la freccia indica frequenza crescente).

Si noti che, come un derivatore, è un sistema passa alto.

- $b_0 = b_1 = \frac{1}{2}$, ovvero un sistema che fa la media degli ultimi due campioni.
Mostriamo i grafici del guadagno e della fase per questo sistema:



e il diagramma di Nyquist

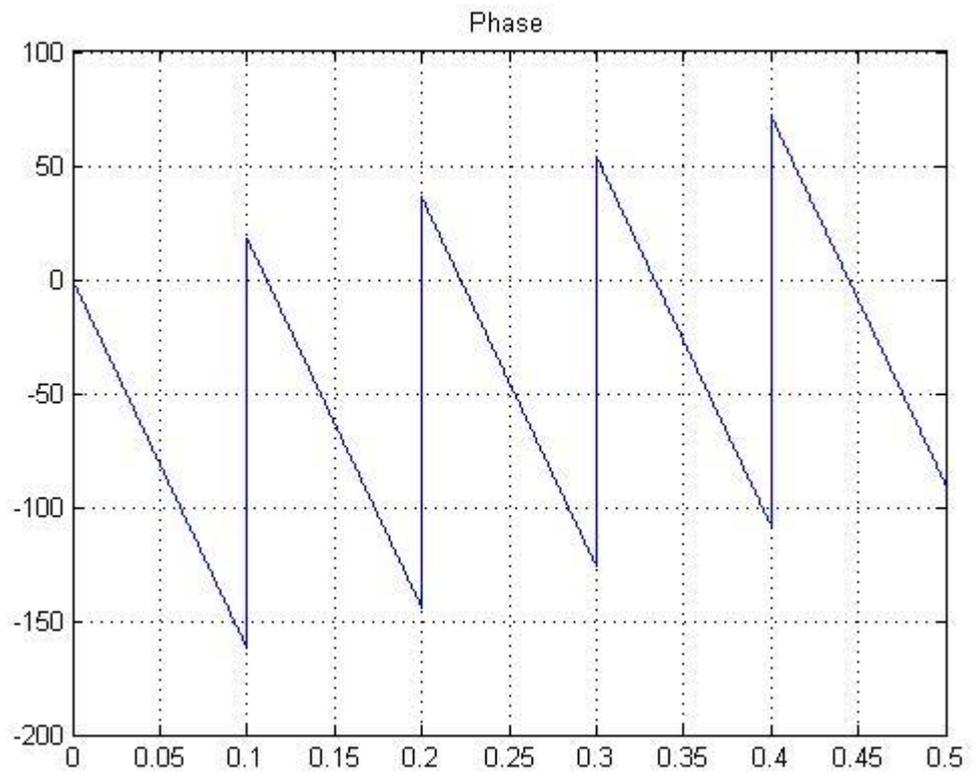
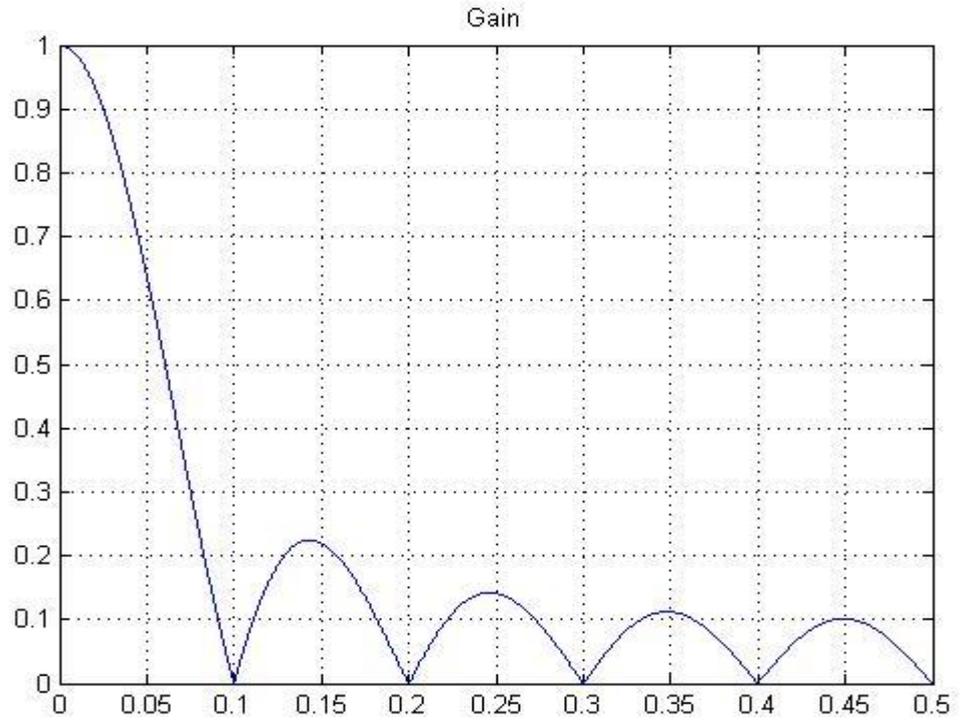


Si vede che si tratta di un sistema passa-basso.

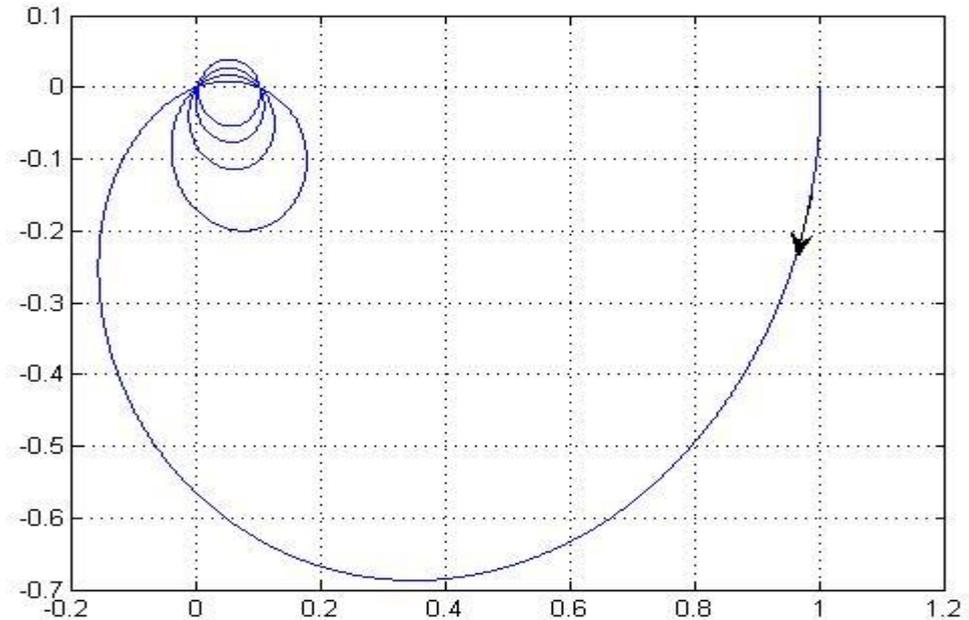
Ci si può domandare cosa accade se la media la facciamo su un numero più elevato di campioni in ingresso, costruendo il sistema

$$(6.56) \quad y_i = \frac{1}{N} \cdot \sum_{k=0}^{N-1} x_{i-k}$$

Consideriamo per esempio un sistema MA di ordine 9, con tutti i 10 coefficienti eguali a 0.1. Ecco i relativi diagrammi:



e Nyquist:



Si noti che è un più “pesante” passa basso, ma presenta delle “risonanze” che potrebbero essere fastidiose.

Due sistemi MA del primo ordine in cascata: MA del secondo ordine

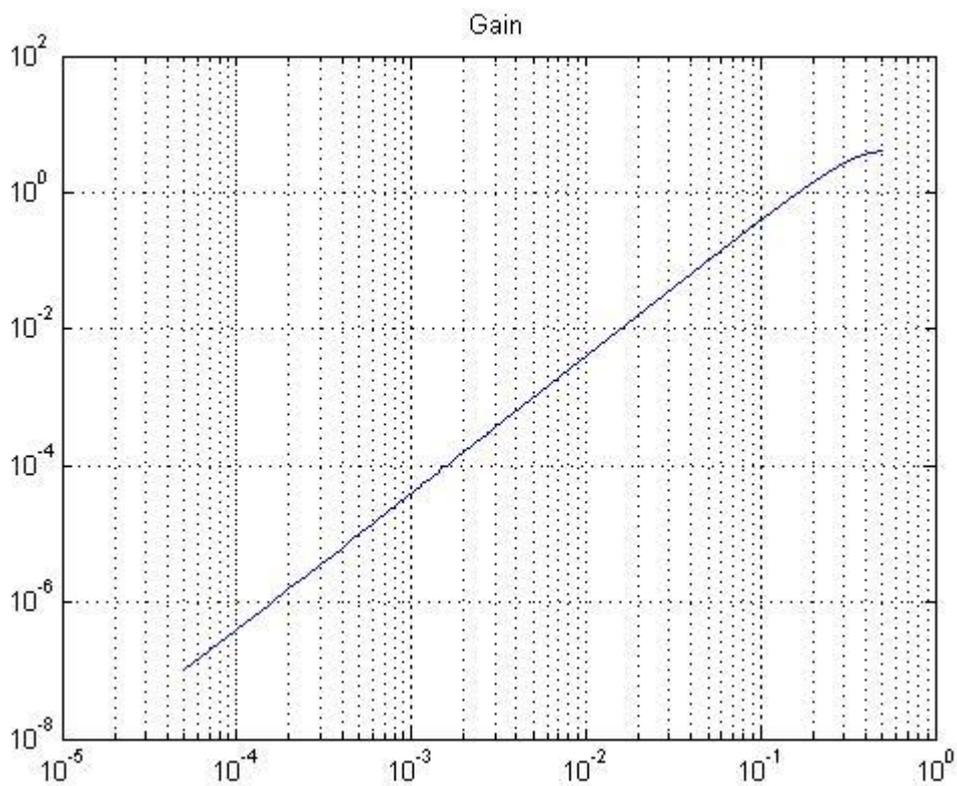
Se poniamo due sistemi in cascata (o “in serie”), la funzione di trasferimento è data dal prodotto delle due funzioni di trasferimento. Se abbiamo due sistemi MA del primo ordine in cascata, con $B^{(1)}(z) = b_0^{(1)} + b_1^{(1)} \cdot z^{-1}$ e $B^{(2)}(z) = b_0^{(2)} + b_1^{(2)} \cdot z^{-1}$, abbiamo

$$(6.57) \quad B(z) = B^{(1)}(z)B^{(2)}(z) = b_0^{(1)}b_0^{(2)} + (b_0^{(1)}b_1^{(2)} + b_1^{(1)}b_0^{(2)})z^{-1} + b_1^{(1)}b_1^{(2)}z^{-2}$$

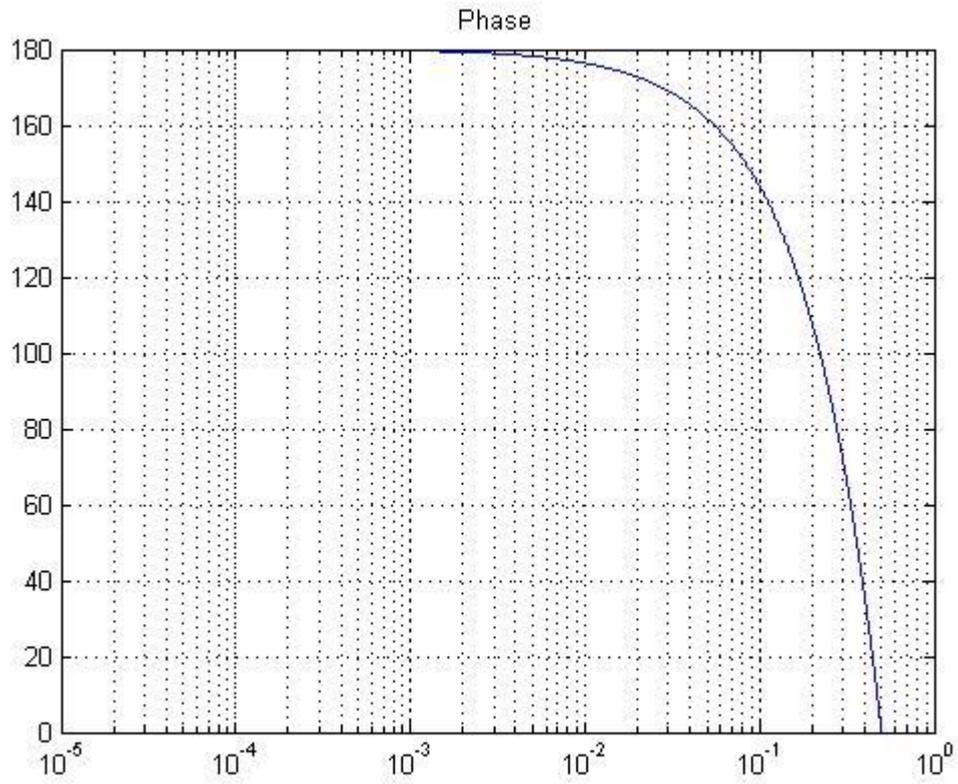
cioè un sistema MA del secondo ordine con coefficienti

$$(6.58) \quad \begin{aligned} b_0 &= b_0^{(1)} b_0^{(2)} \\ b_1 &= b_0^{(1)} b_1^{(2)} + b_1^{(1)} b_0^{(2)} \\ b_2 &= b_1^{(1)} b_1^{(2)} \end{aligned}$$

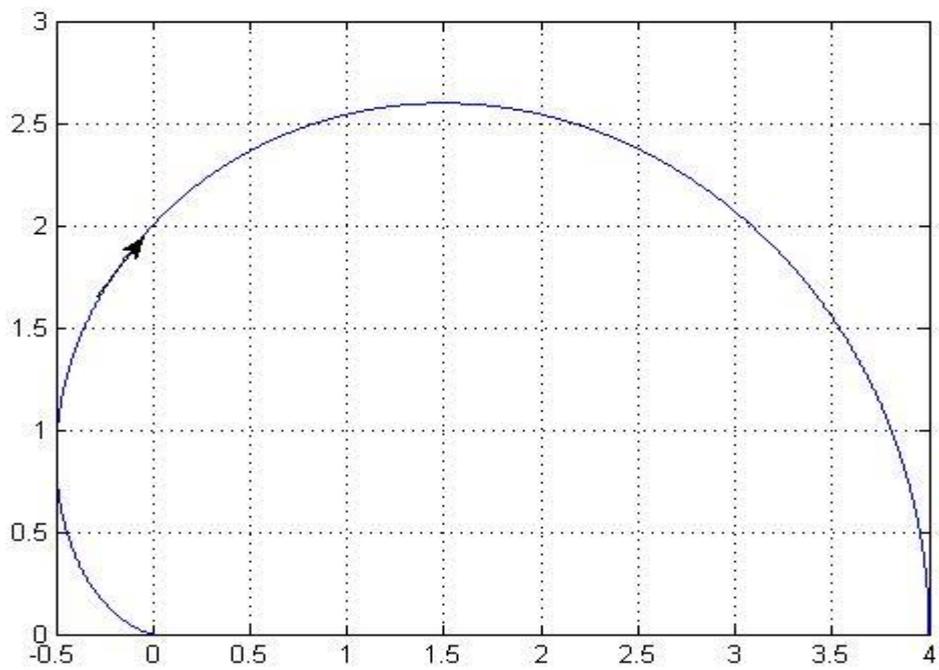
Se per esempio mettiamo in cascata due sistemi con $b_0 = -b_1 = 1$, cioè che eseguono la differenza prima, abbiamo un sistema con $b_0 = b_2 = 1$, $b_1 = -2$; questo è un sistema che esegue la differenza seconda (analoga alla derivata seconda del continuo). Abbiamo:



Si noti la pendenza di due decadi per ogni decade di frequenza, doppia che per la differenza prima.



e Nyquist



Sistema AR del primo ordine (reale)

Consideriamo il sistema in cui nell'equazione (6.48) , $n=1$, $m=0$; abbiamo

$$(6.59) \quad y_i = b_0 \cdot x_i - a_1 \cdot y_{i-1}$$

In questo caso la risposta impulsiva è facilmente calcolabile come

$$(6.60) \quad \begin{aligned} x_0 &= b_0 \\ x_1 &= -b_0 \cdot a_1 \\ x_2 &= b_0 \cdot a_1^2 \\ x_3 &= -b_0 \cdot a_1^3 \\ &\dots \\ x_i &= b_0 \cdot (-a_1)^i \\ &\dots \end{aligned}$$

Si vede immediatamente che se $|a_1| > 1$ allora il sistema non è stabile (la risposta impulsiva ha un andamento esponenziale crescente).

Se $|a_1| \leq 1$ il sistema è stabile e, se $a_1 < 0$, la risposta impulsiva è esponenziale decrescente, altrimenti è in modulo esponenziale decrescente, ma con segni alternativamente positivi e negativi. L'andamento esponenziale ha decadimento τ dato da

$$(6.61) \quad \tau = -\frac{1}{\ln |a_1|}$$

che, per valori di $|a_1|$ molto vicini a 1, si può approssimare con

$$(6.62) \quad \tau \approx \frac{1}{1 - |a_1|}$$

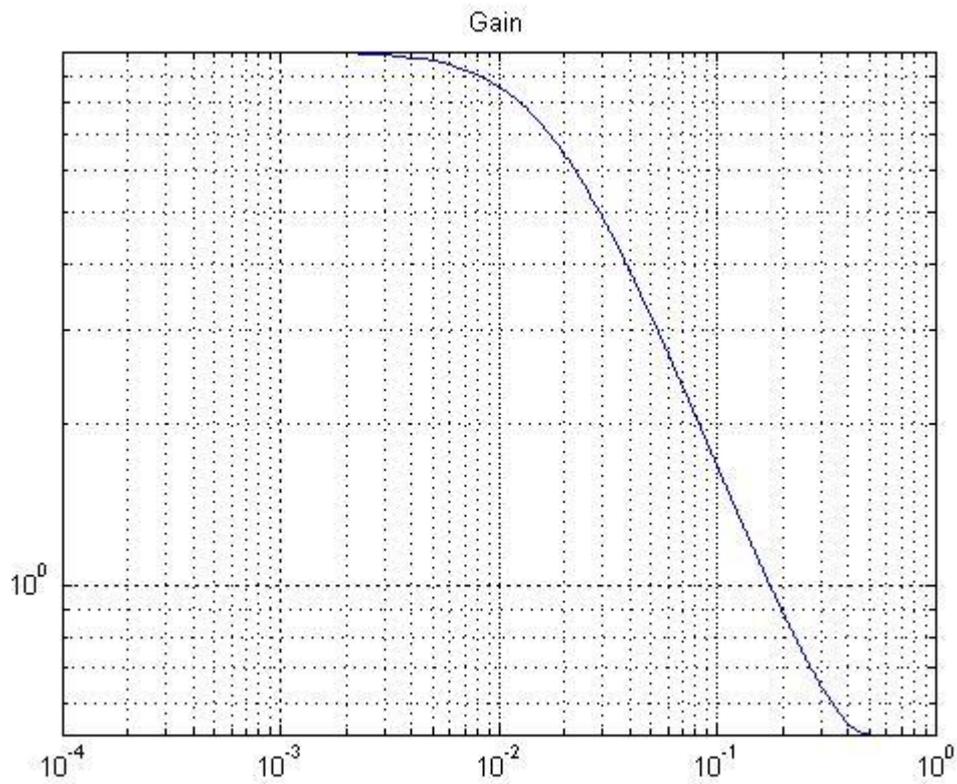
Se $a_1 = -1$, la risposta impulsiva è un gradino di ampiezza b_0 .

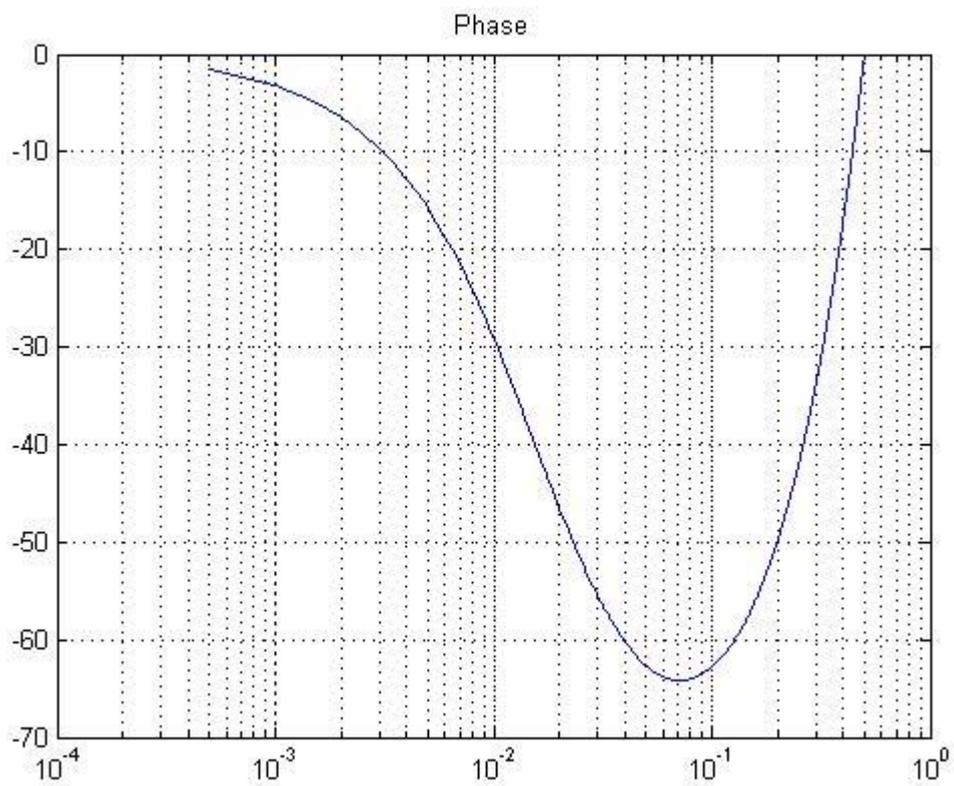
Vediamo ora la risposta in frequenza. Si ha

$$(6.63) \quad \mathcal{F}(\Omega) = F(e^{j\Omega}) = \frac{b_0}{1 + a_1 \cdot e^{-j\Omega}}$$

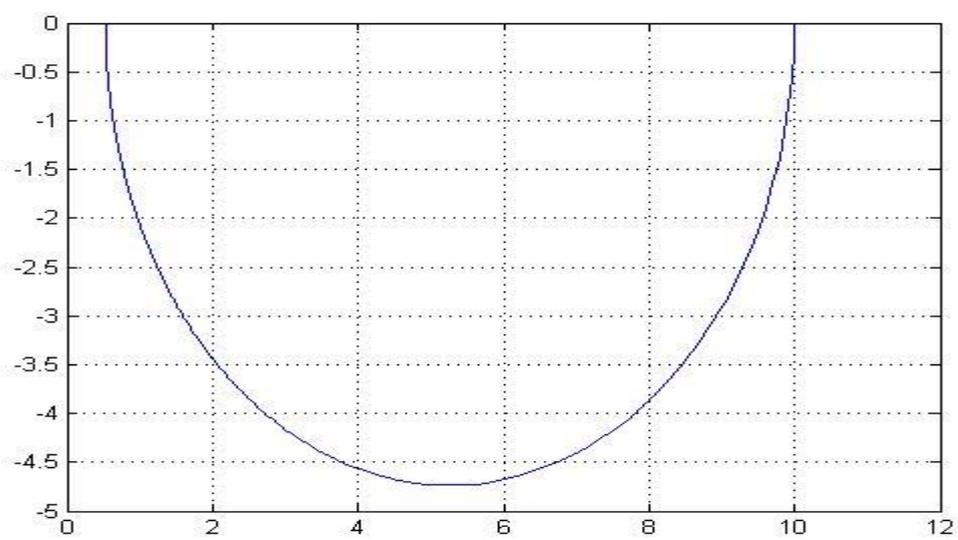
Grafichiamo alcuni casi:

- $b_0 = 1, a_1 = -0.9$

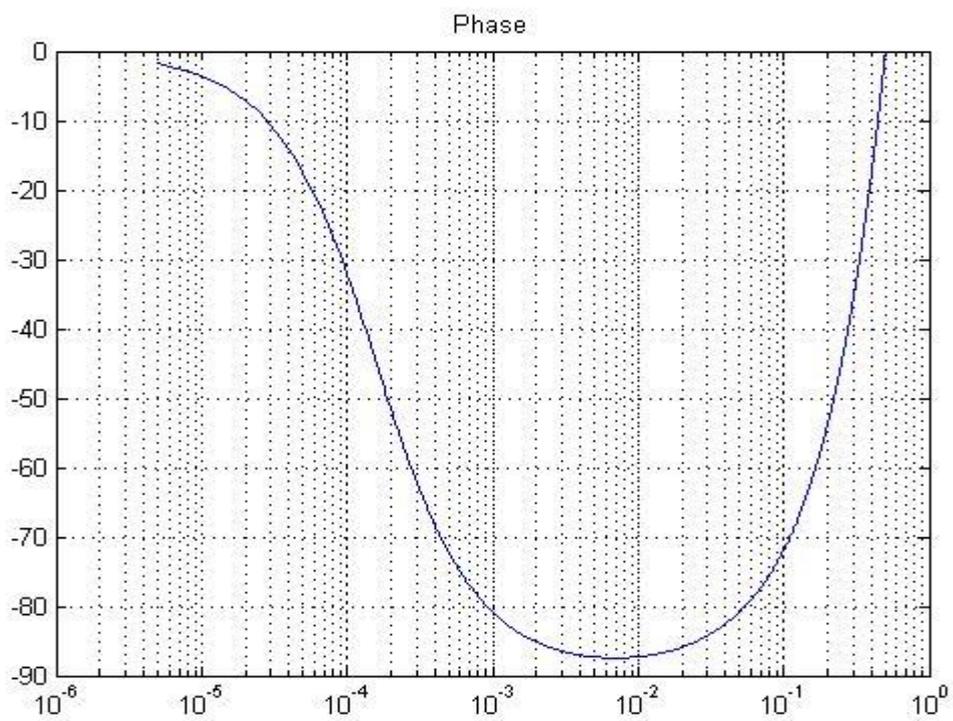
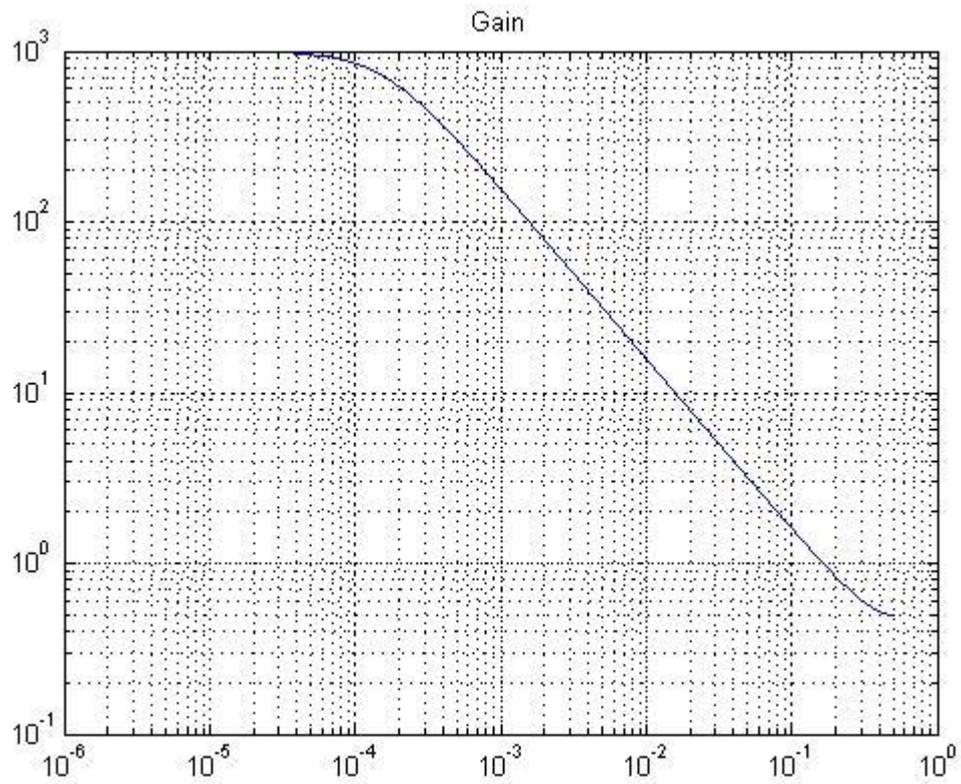


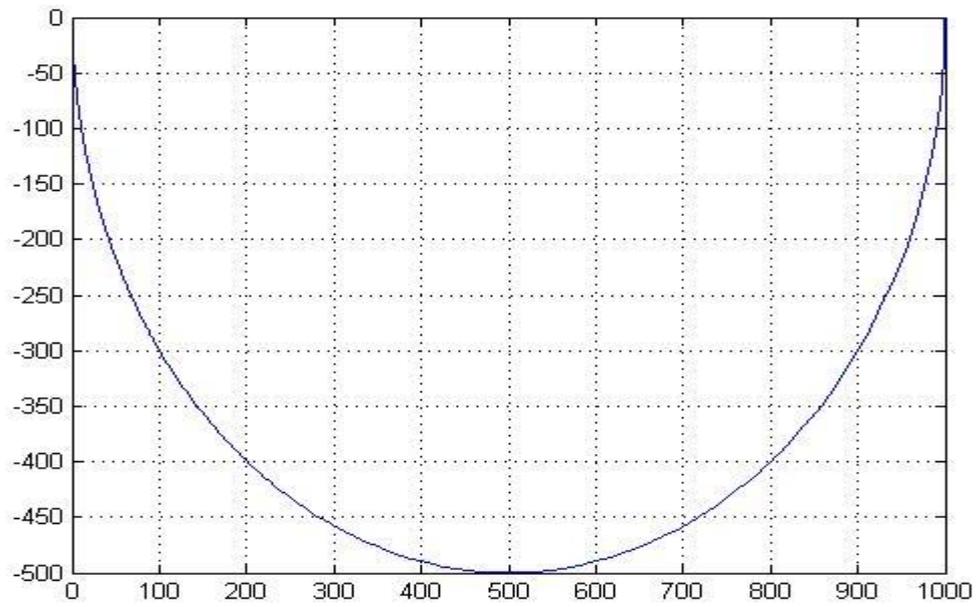


E il diagramma di Nyquist

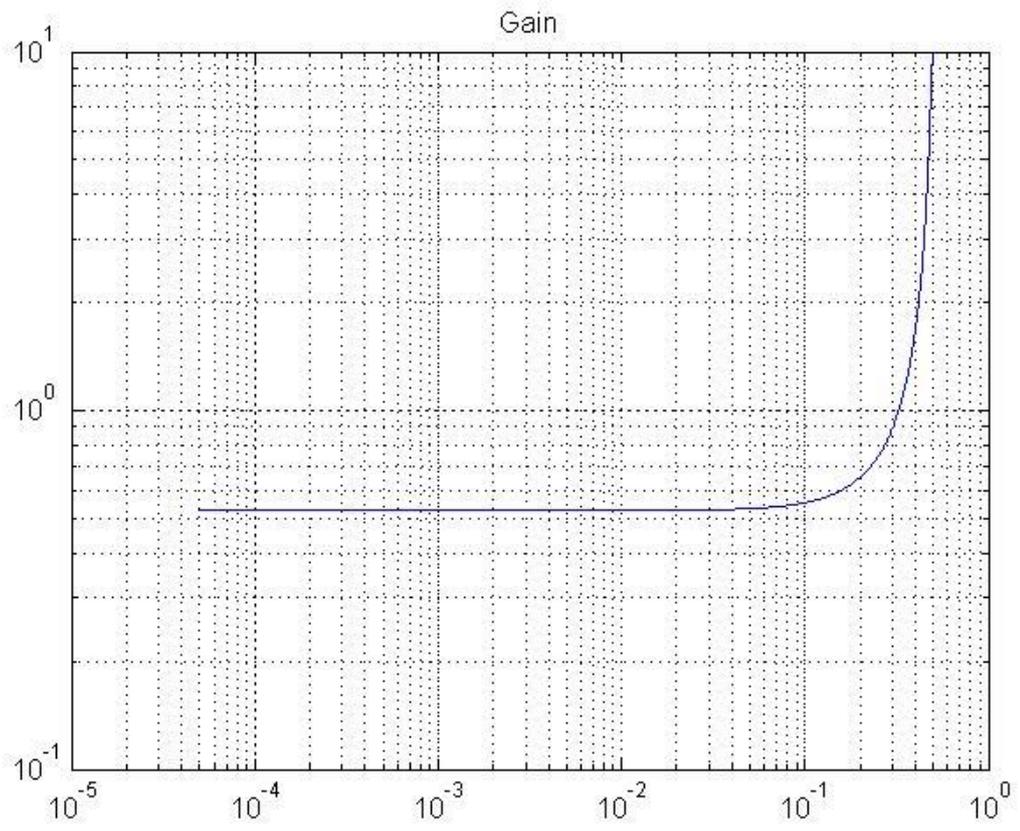


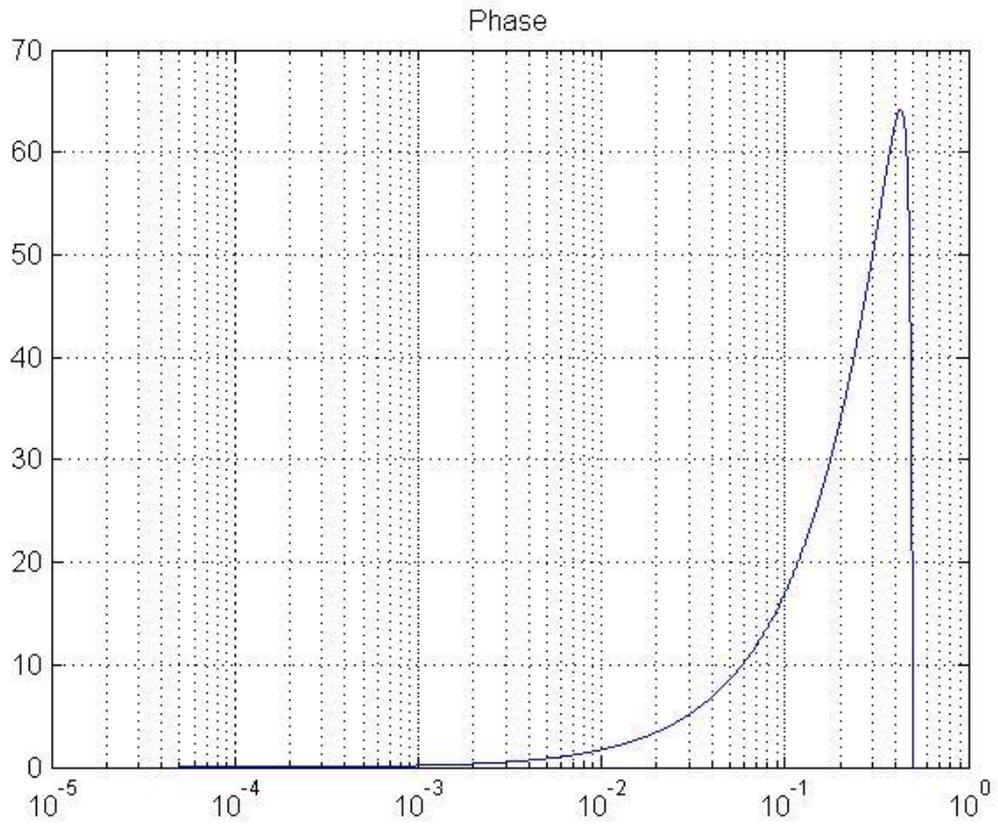
- $b_0 = 1, a_1 = -0.999$



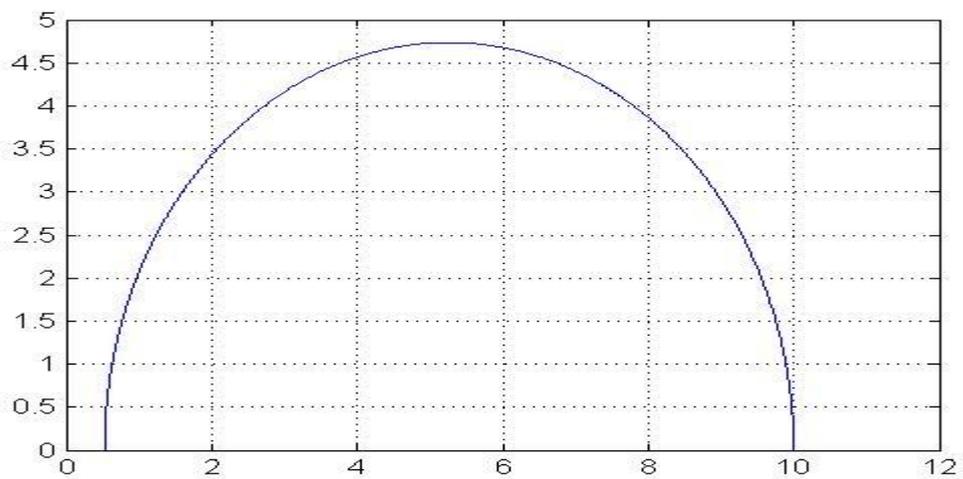


- $b_0 = 1, a_1 = 0.9$. In questo caso si ha



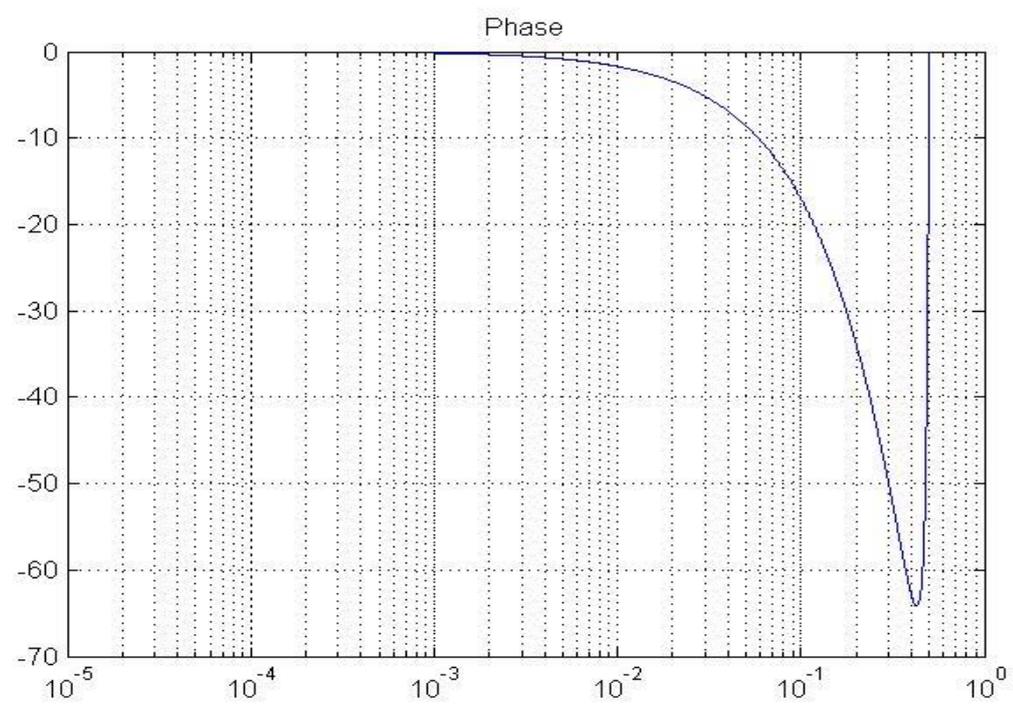
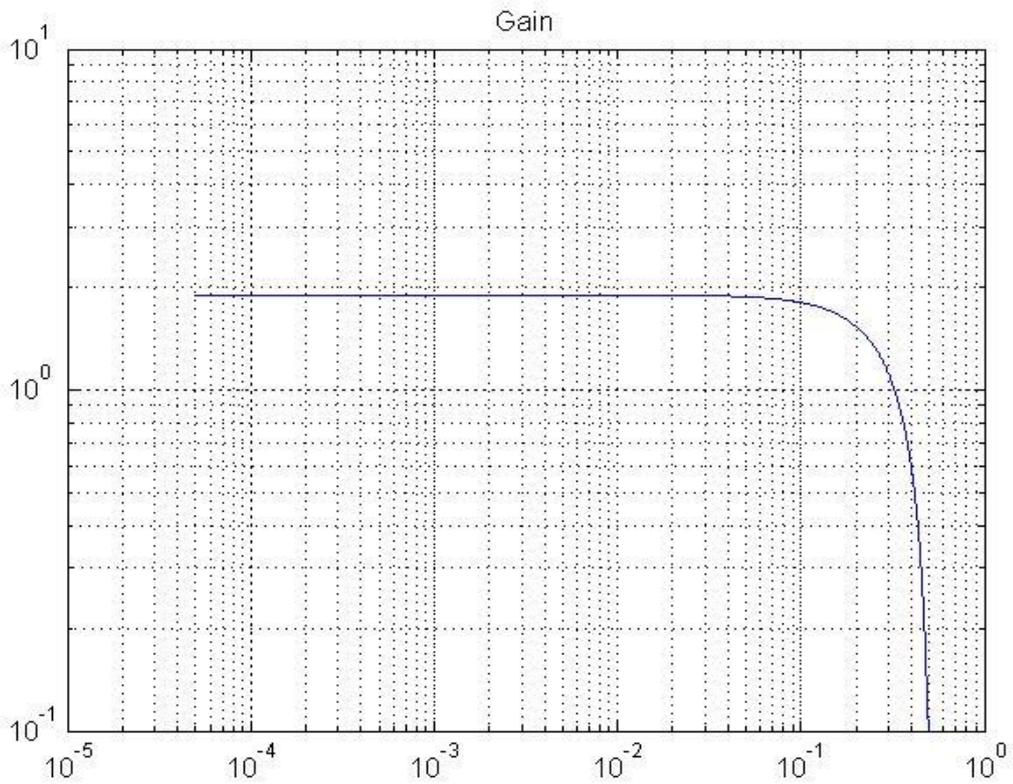


e Nyquist

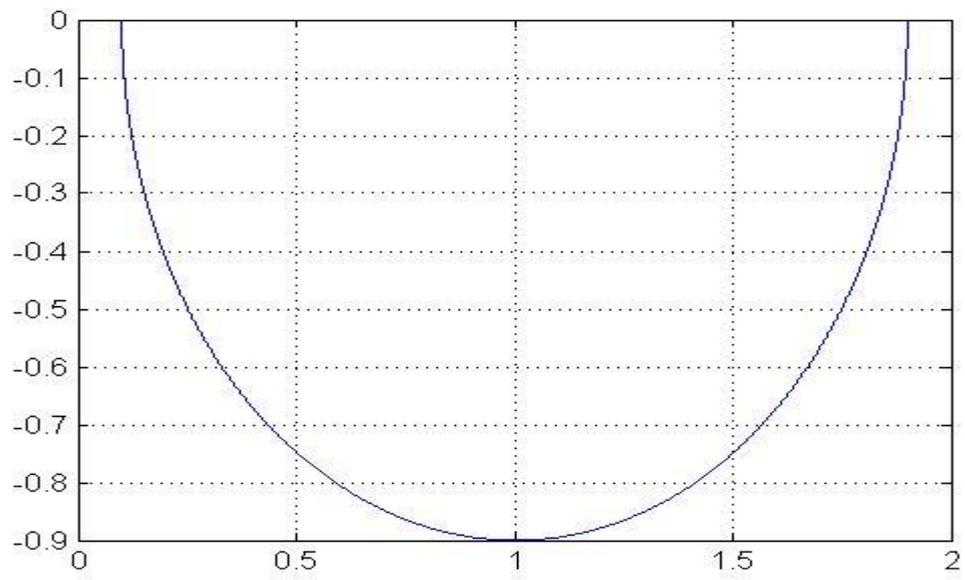


- Vediamo ora cosa accade se consideriamo l'inverso del sistema precedente, in cui semplicemente nella funzione di trasferimento z si è invertito il numeratore

col denominatore. Il sistema era $b_0 = 1$, $a_1 = 0.9$; il suo inverso è un sistema MA che ha $b_0 = 1$, $b_1 = -0.9$. Ecco i grafici:



e Nyquist



Sistema AR del primo ordine (complesso)

Se nell'equazione (6.59) il coefficiente a_1 è complesso, formalmente le cose non cambiano: la risposta impulsiva può essere espressa sempre come in equazione (6.60), ma ora, come vedremo, siamo in presenza di un'oscillazione smorzata.

Sia

$$(6.64) \quad w = -a_1 = r \cdot e^{j\varphi}$$

a parte il coefficiente di guadagno b_0 , la risposta impulsiva mostra tutte le potenze di w , da 0 a infinito; il k -esimo termine è dato da

$$(6.65) \quad x_k = b_0 \cdot w^k = b_0 \cdot (r^k \cdot e^{j \cdot k \cdot \varphi})$$

Notiamo che:

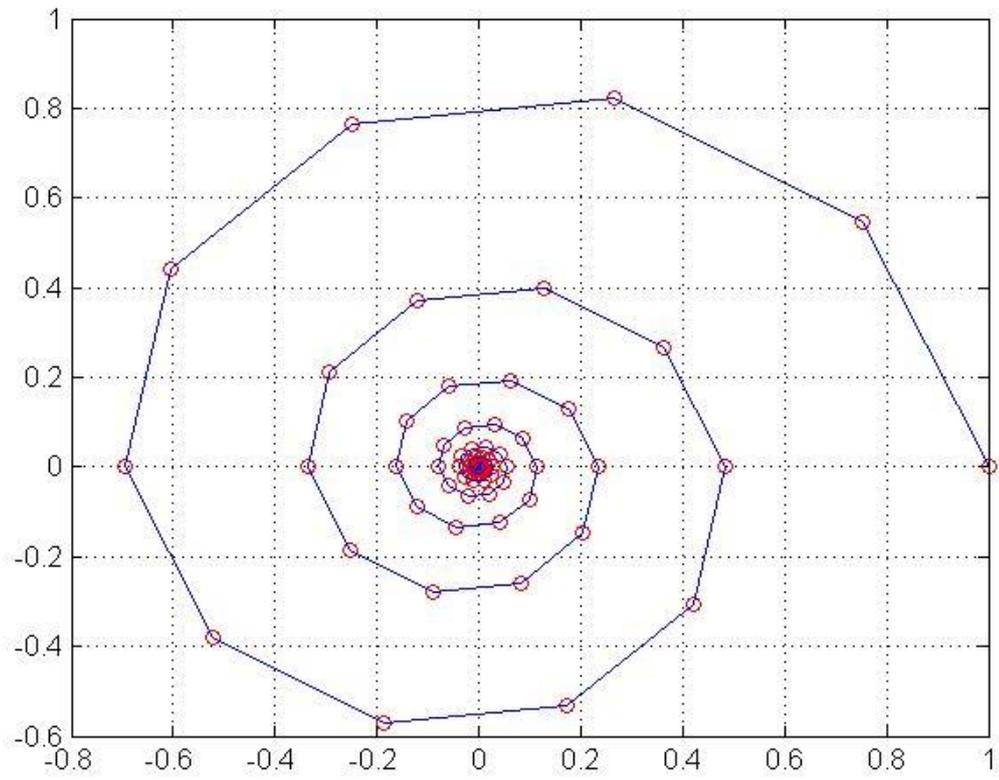
- il sistema è stabile se $r < 1$
- in tale ipotesi il valore assoluto decresce esponenzialmente
- la fase di w^k (e quindi di x_k) cresce proporzionalmente a k , quindi il vettore che rappresenta nel piano di Gauss il numero complesso x_k ruota intorno all'origine con velocità proporzionale a φ , riducendosi proporzionalmente di ampiezza. La frequenza di rotazione è data da

$$(6.66) \quad \nu_0 = \frac{\varphi}{2\pi}$$

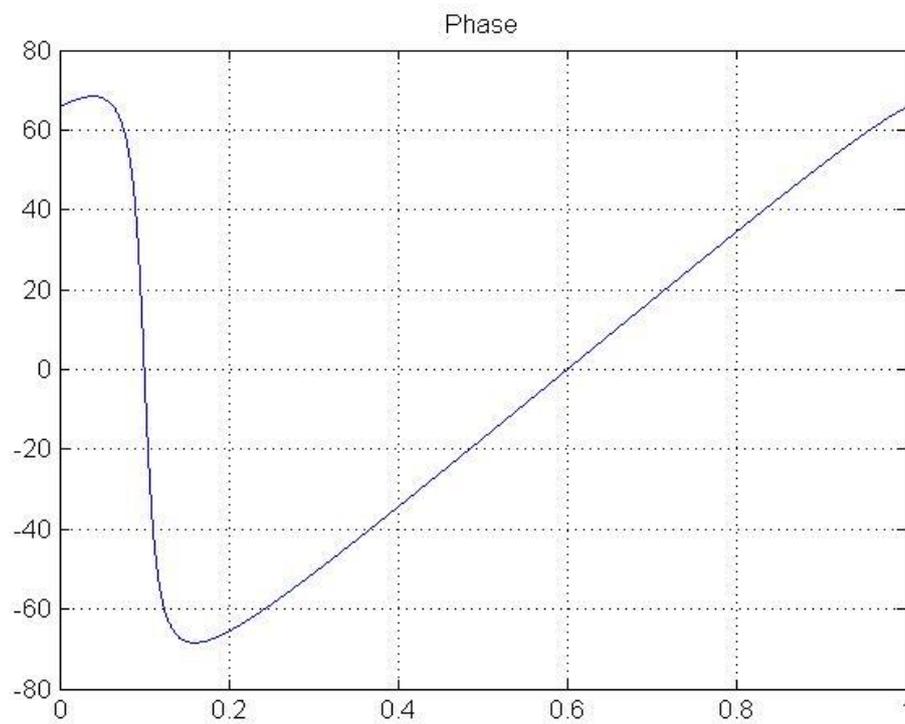
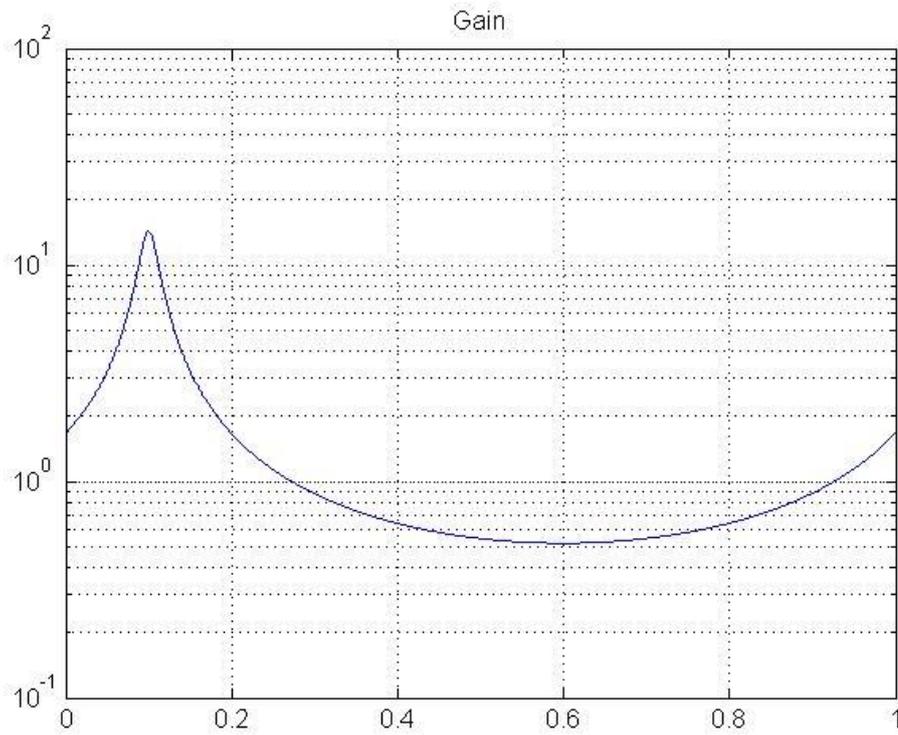
e il tempo di decadimento τ è dato da (vedi equazione (6.61))

$$(6.67) \quad \tau = -\frac{1}{\ln |w|} = -\frac{1}{\ln r}$$

Consideriamo il caso di $b_0 = 1$, $r = 0.95$, $\varphi = 36^\circ$. Ecco l'evoluzione nel piano di Gauss

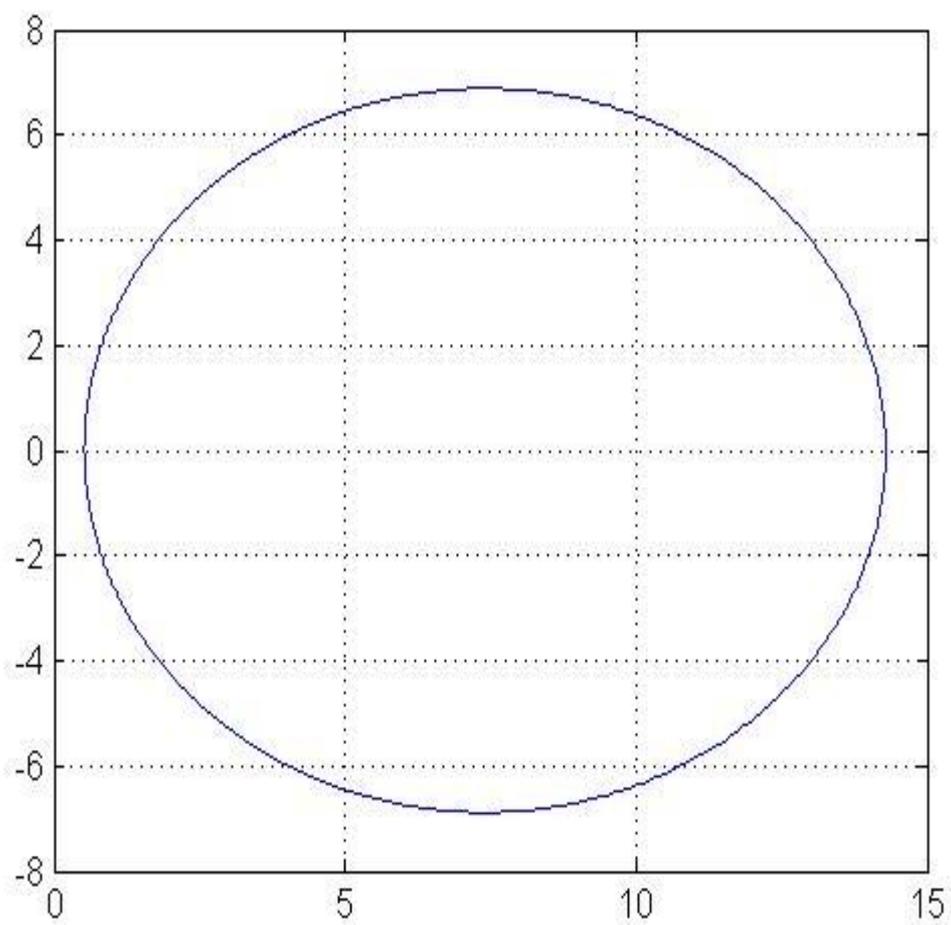


La risposta in frequenza è data da



Si noti la presenza di una risonanza a frequenza $0.1 \left(\frac{36}{360} = 0.1 \right)$.

Il diagramma di Nyquist è



Sistema AR del secondo ordine

Consideriamo il sistema in cui nell'equazione (6.48) , $n=2$, $m=0$; abbiamo

$$(6.68) \quad y_i = b_0 \cdot x_i - a_1 \cdot y_{i-1} - a_2 \cdot y_{i-2}$$

Il denominatore della funzione di trasferimento è

$$(6.69) \quad A(z) = 1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} = (1 - p_1 \cdot z^{-1}) \cdot (1 - p_2 \cdot z^{-1})$$

dove p_1 e p_2 sono i due poli (soluzioni della $A(z)=0$).

Se i due poli sono complessi coniugati, li possiamo porre come

$$(6.70) \quad p_{1,2} = r \cdot e^{\pm j\theta}$$

e quindi

$$(6.71) \quad \begin{aligned} A(z) &= 1 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2} = 1 - (p_1 + p_2) \cdot z^{-1} + (p_1 \cdot p_2) \cdot z^{-2} = \\ &= 1 - 2 \cdot r \cdot \cos \theta \cdot z^{-1} + r^2 \cdot z^{-2} \end{aligned}$$

Se $r < 1$ il sistema è stabile e, se $a_1 < 0$, la risposta impulsiva è una sinusoide smorzata esponenzialmente

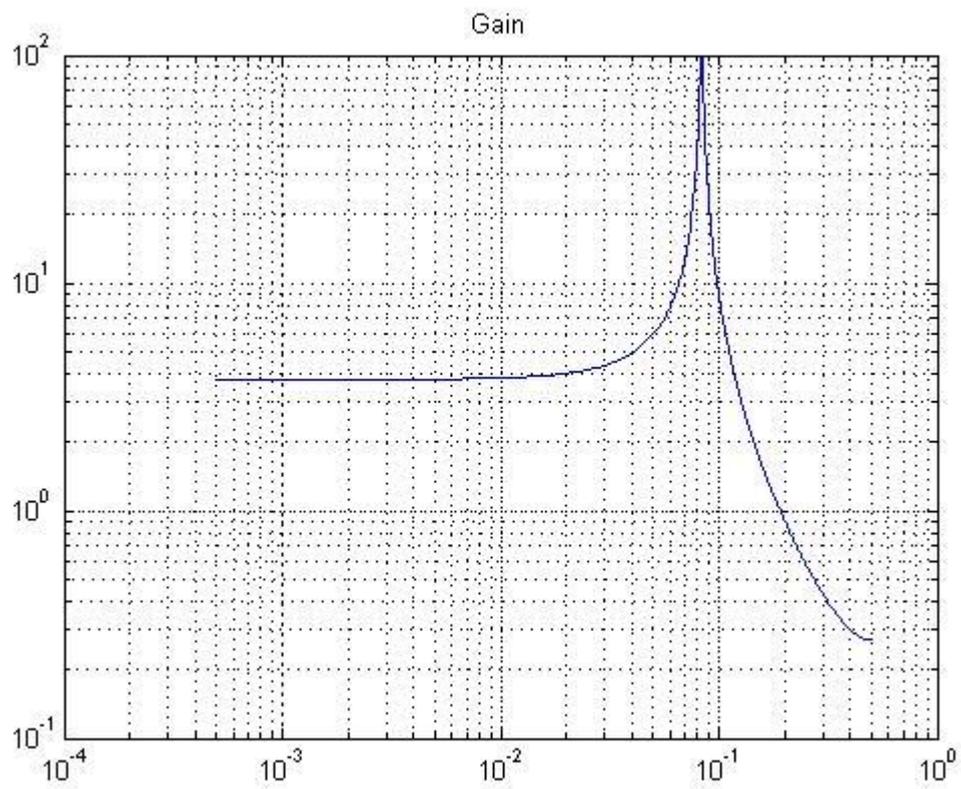
$$(6.72) \quad \eta_i = \frac{b_0}{2} \cdot (p_1^i + p_2^i) = b^0 \cdot r^i \cdot \cos(i \cdot \theta)$$

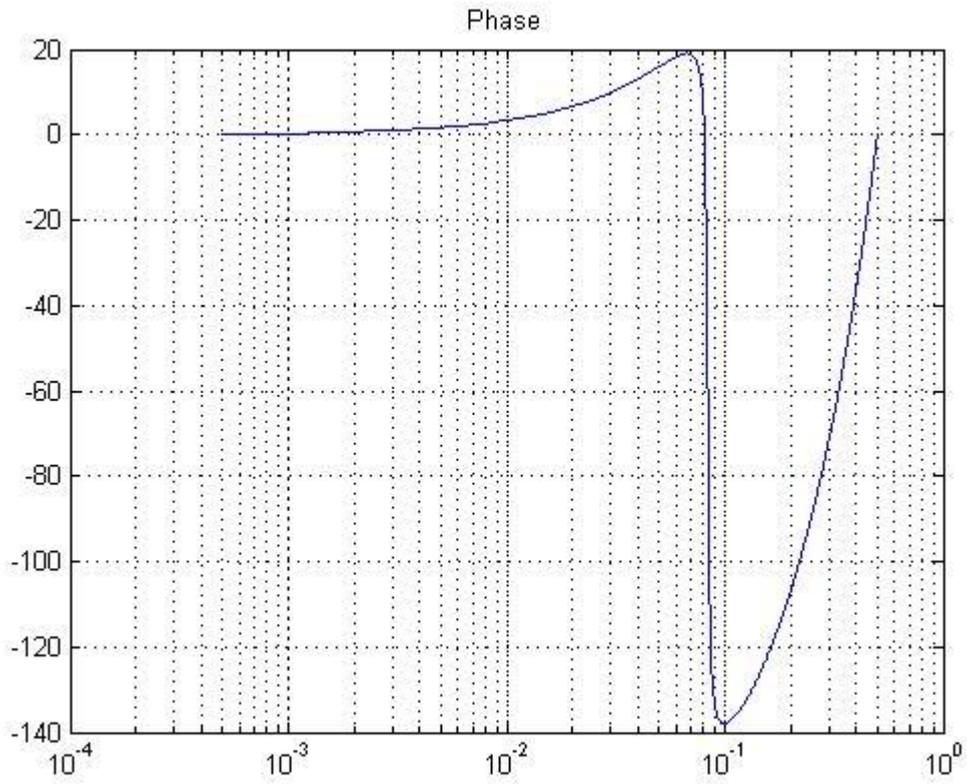
L'andamento esponenziale ha tempo di decadimento τ dato da

(6.73)

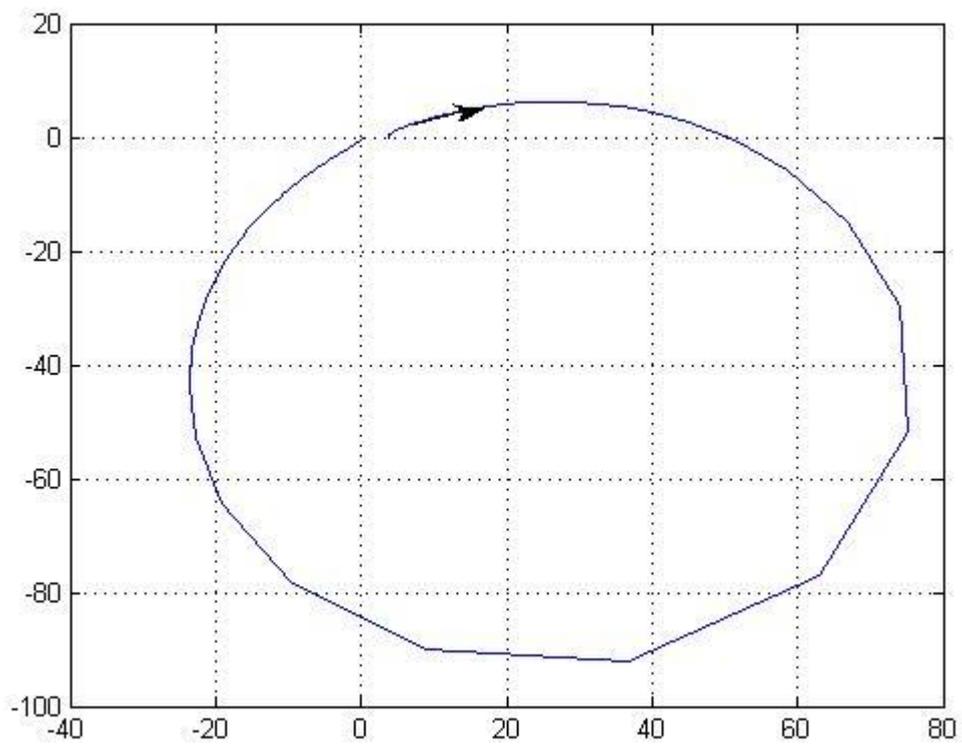
$$\tau = -\frac{1}{\ln r}$$

Grafichiamo il caso in cui $b_0 = 1$, $r = 0.99$, $\theta = \frac{\pi}{6}$. Abbiamo





e il diagramma di Nyquist



Semplici sistemi ARMA

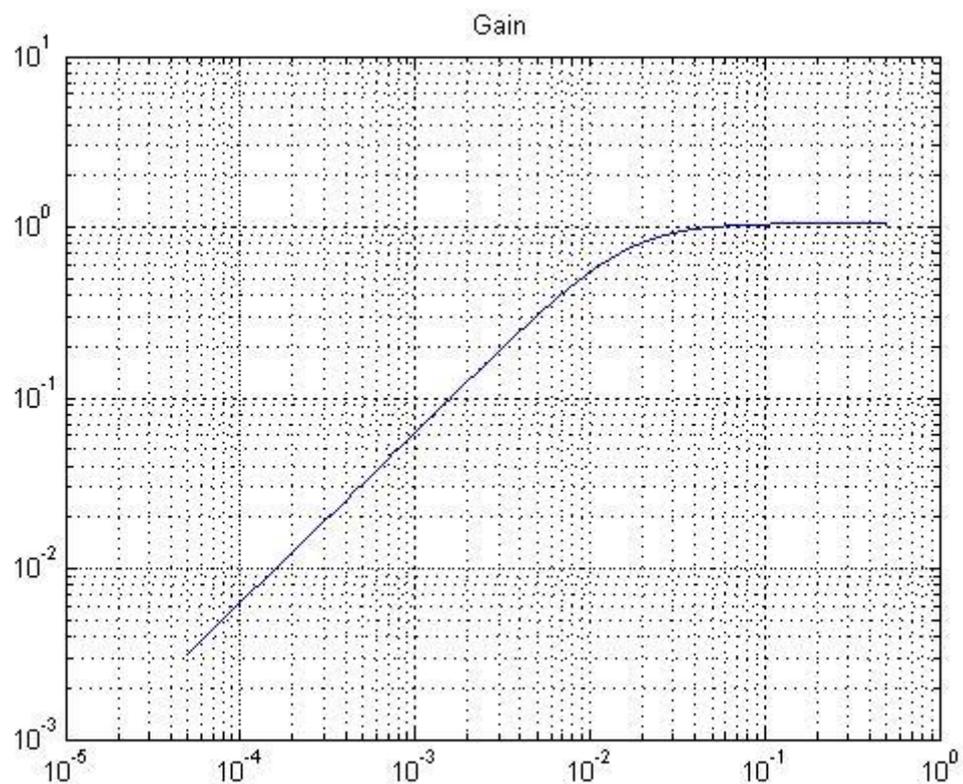
Come esempi di sistemi di questo tipo prendiamo i seguenti:

- sistema passa-alto (ARMA[1,1]), con funzione di trasferimento

$$(6.74) \quad F(z) = \frac{1 - z^{-1}}{1 - w \cdot z^{-1}}$$

derivato dal passa-basso.

Per $w=0.9$ si ha



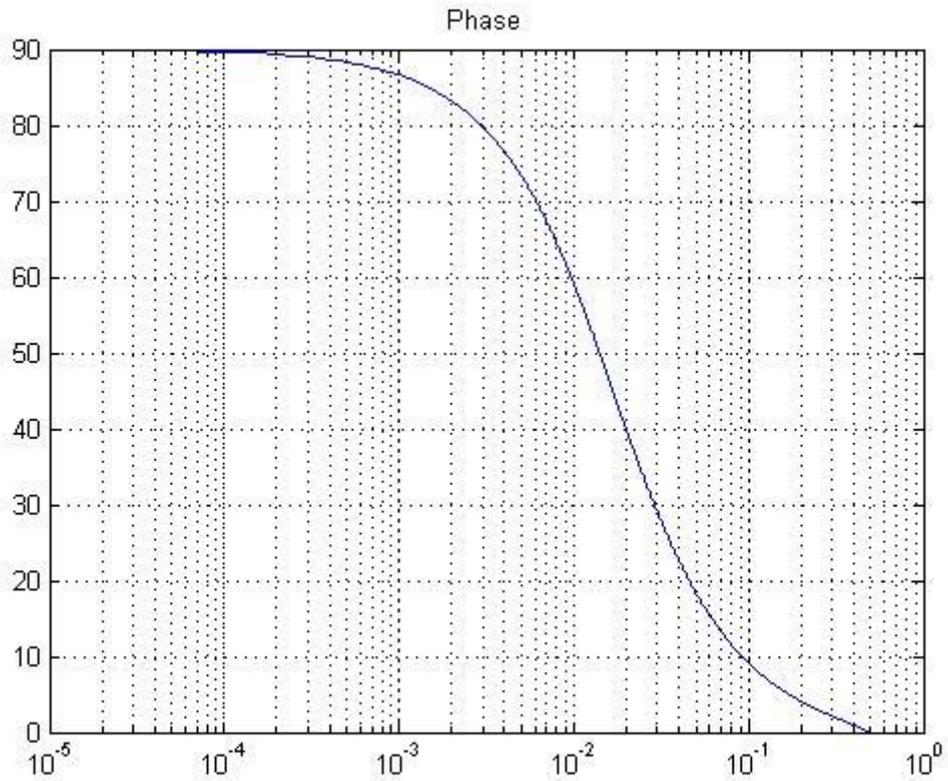
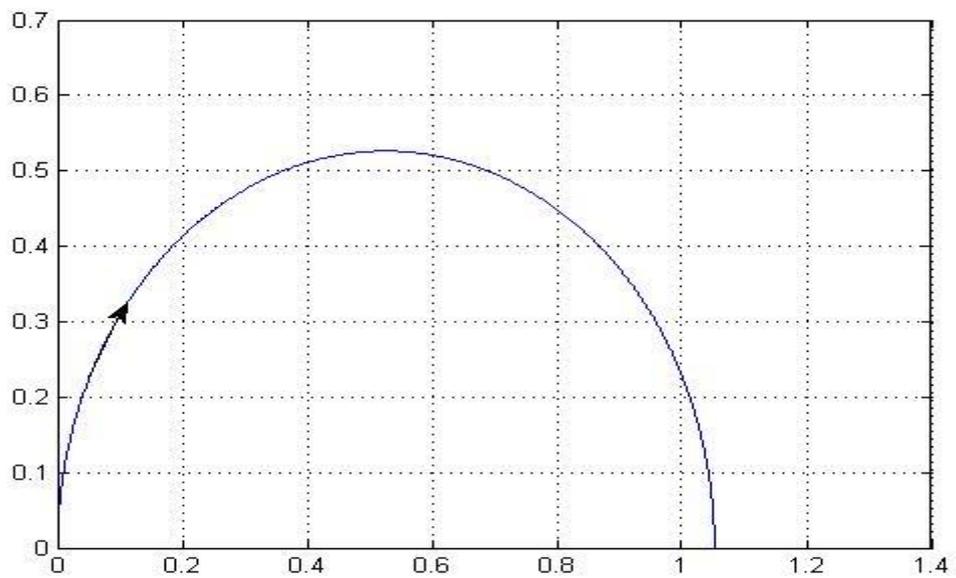


Diagramma di Nyquist

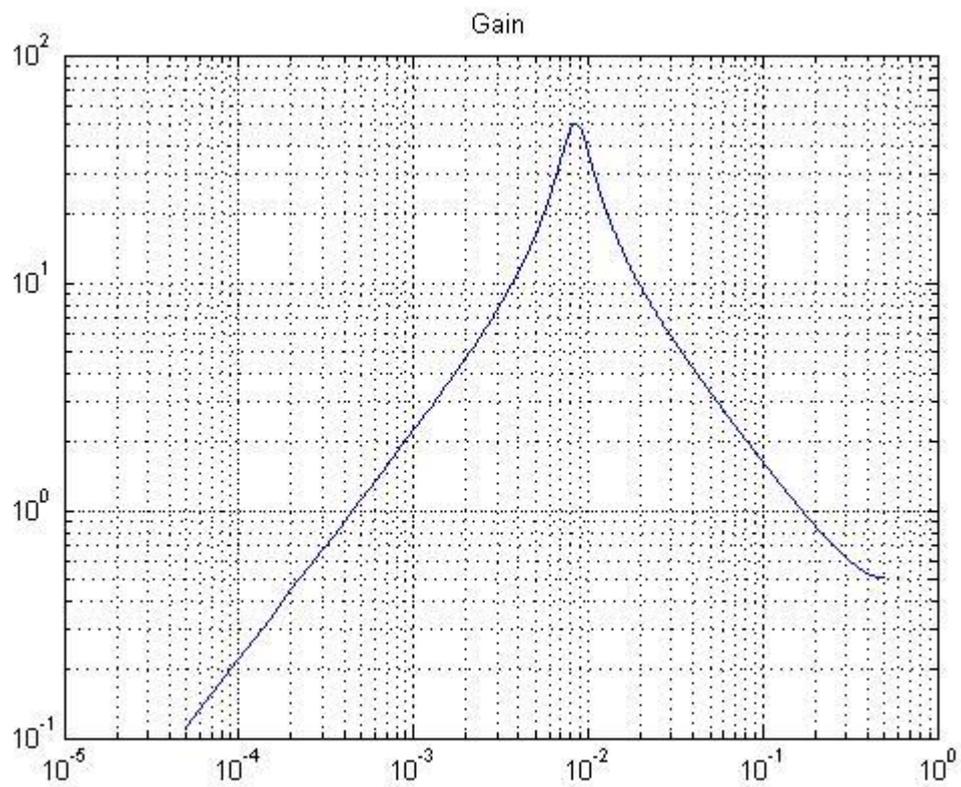


- risonanza passa-banda (ARMA[2,1]), con funzione di trasferimento

$$(6.75) \quad F(z) = \frac{1 - z^{-1}}{1 - 2r \cos \theta \cdot z^{-1} + r^2 \cdot z^{-2}}$$

derivato dall'AR del secondo ordine.

Per $r = 0.99$, $\theta = \frac{\pi}{60}$ si ha



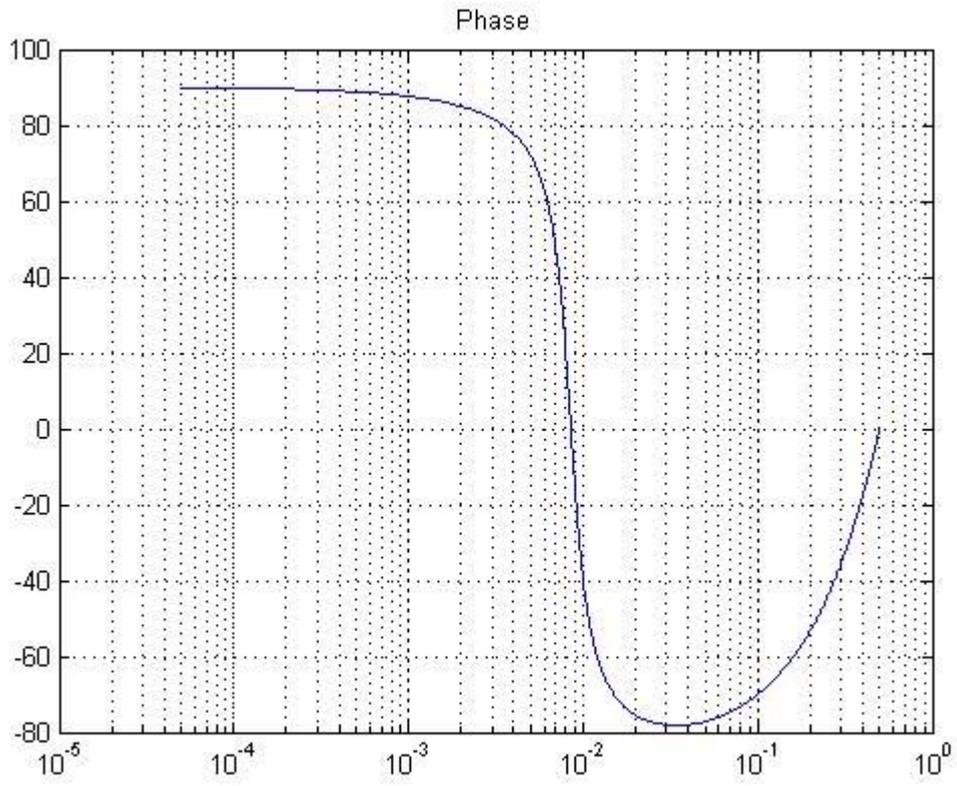
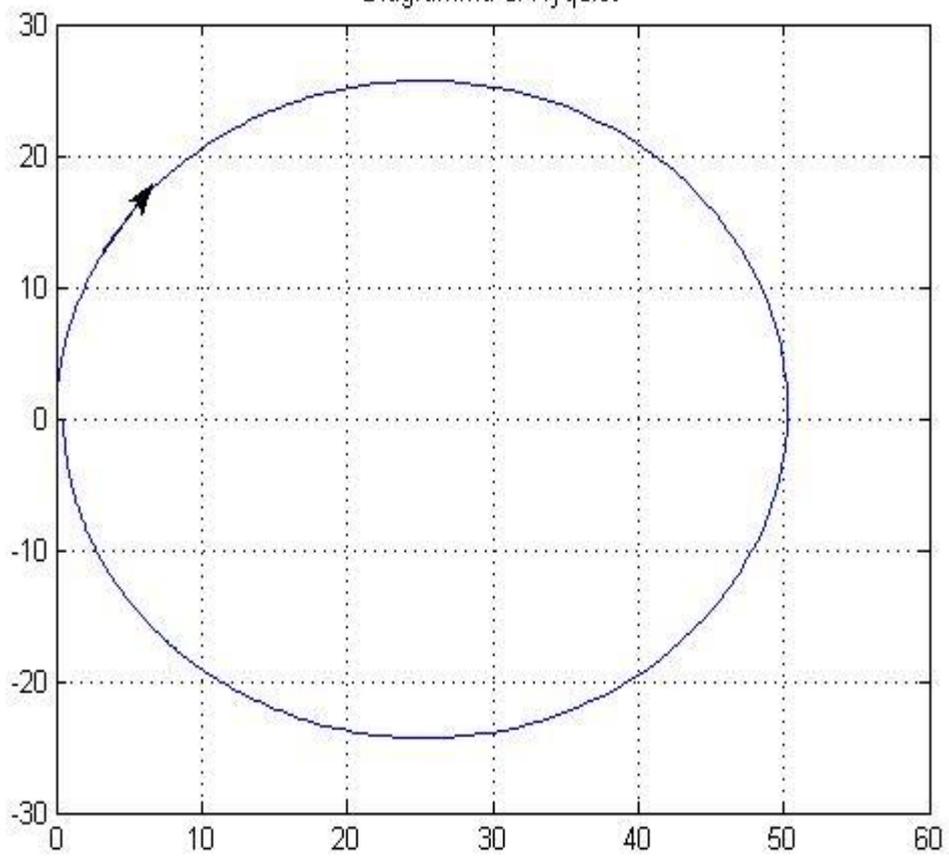


Diagramma di Nyquist



Sistemi non-lineari - i sistemi di Volterra

Facciamo un breve cenno ai sistemi non-lineari, una categoria così vasta che non esistono metodi di analisi generali (a parte cose piuttosto generiche).

Ci limiteremo a quelli che sono considerati il primo passo della non-linearità: i sistemi di Volterra del secondo ordine (i sistemi di Volterra del primo ordine sono semplicemente i sistemi lineari discreti). Essi sono descritti dall'equazione

$$(6.76) \quad y_i = \sum_{k=0}^{\infty} h_k \cdot x_{i-k} + \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} q_{kl} \cdot x_{i-k} \cdot x_{i-l}$$

come si vede, oltre al membro lineare presente già nella (6.9), ce ne è un altro quadratico.

Si trova che nel dominio trasformato si ha

$$(6.77) \quad Y(\omega) = H(\omega) \cdot X(\omega) + \frac{1}{2} \cdot \int_{-\infty}^{\infty} Q(\omega_1, \omega - \omega_1) \cdot X(\omega_1) \cdot X(\omega - \omega_1) \cdot d\omega_1$$

7 - Segnali aleatori

Live Script

Introduzione ai processi stocastici

Definizione di processo stocastico

Ricordiamo la nostra definizione di variabile casuale:

“Una variabile casuale è una variabile (reale o complessa) al cui valore è associata una distribuzione di probabilità.”

In analogia, possiamo dare la seguente definizione di processo stocastico:

Un processo stocastico è un insieme di funzioni del tempo alle quali è associata una distribuzione di probabilità.

Ciascuna delle funzioni del tempo viene chiamata **realizzazione** del processo stocastico. Mentre il processo stocastico è il modello generale, quello che si può osservare sperimentalmente è soltanto una realizzazione.

Se le funzioni del tempo sono continue, parliamo di **processo stocastico continuo**, se sono discrete di **processo stocastico discreto**. Nel seguito svilupperemo il formalismo per lo più per il caso continuo, ma analoghi sviluppi si hanno nel caso discreto (dove, ricordiamo, la variabile t è discreta).

Per un processo stocastico $X(t)$, possiamo definire la funzione di distribuzione del primo ordine

$$(7.1) \quad F(x;t) = P\{X(t) \leq x\}$$

da cui la densità

$$(7.2) \quad f(x;t) = \frac{\partial F(x;t)}{\partial x}$$

Abbiamo quindi il valor medio (o valore atteso)

$$(7.3) \quad \mu(t) = E[X(t)] = \int_{-\infty}^{\infty} x \cdot f(x;t) \cdot dx$$

e la varianza

$$(7.4) \quad \sigma^2(t) = E\left[\left(X(t) - \mu(t)\right)^2\right] = \int_{-\infty}^{\infty} (x - \mu(t))^2 \cdot f(x;t) \cdot dx$$

Per un dato valore del tempo t , un processo stocastico si riduce a una variabile casuale.

Si definiscono inoltre, per qualsiasi $n > 0$, le funzioni di distribuzione di ordine n come

$$(7.5) \quad F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\}$$

e le relative densità

$$(7.6) \quad f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}$$

Un processo stocastico è statisticamente definito se sono definite tutte le funzioni di distribuzione (o di densità) per tutti gli ordini.

Dato un processo stocastico, possiamo crearne uno nuovo facendolo passare attraverso un opportuno sistema. Il nuovo processo ha in genere caratteristiche differenti da quello di partenza.

Funzioni del secondo ordine

Consideriamo le funzioni del secondo ordine,

$$(7.7) \quad F(x_1, x_2; t_1, t_2) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2\}$$

e

$$(7.8) \quad f(x_1, x_2; t_1, t_2) = \frac{\partial^2 F(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2}$$

Si noti che

$$(7.9) \quad F(x_1, \infty; t_1, t_2) = F(x_1; t_1)$$

e

$$(7.10) \quad f(x_1; t_1) = \int_{-\infty}^{\infty} f(x_1, x_2; t_1, t_2) \cdot dx_2$$

Possiamo calcolare la densità condizionale come

$$(7.11) \quad f(x_1; t_1 | X(t) = x_2) = \frac{f(x_1, x_2; t_1, t_2)}{f(x_2; t_2)}$$

Definiamo **autocorrelazione** la funzione

$$(7.12) \quad R(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \cdot f(x_1, x_2; t_1, t_2) \cdot dx_1 \cdot dx_2$$

e **autocovarianza** la funzione

$$(7.13) \quad \begin{aligned} C(t_1, t_2) &= E[(X(t_1) - \mu(t_1)) \cdot (X(t_2) - \mu(t_2))] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \mu(t_1)) \cdot (x_2 - \mu(t_2)) \cdot f(x_1, x_2; t_1, t_2) \cdot dx_1 \cdot dx_2 = \\ &= R(t_1, t_2) - \mu(t_1) \cdot \mu(t_2) \end{aligned}$$

la varianza si può ricavare come

$$(7.14) \quad \sigma^2(t) = C(t, t) = R(t, t) - \mu^2(t)$$

Se il processo è complesso (cioè se sono complesse le funzioni del tempo), allora

$$(7.15) \quad R(t_1, t_2) = E[X(t_1) \cdot X^*(t_2)]$$

e

$$(7.16) \quad C(t_1, t_2) = E[(X(t_1) - \mu(t_1)) \cdot (X^*(t_2) - \mu^*(t_2))]$$

In questo caso C ed R sono funzioni complesse. Si noti che è una generalizzazione del caso reale.

Dall'autocorrelazione $R(t_1, t_2) = R(t_1, t_1 + \tau) = R(t, t + \tau)$ possiamo ricavare lo **spettro di potenza** (spesso chiamato semplicemente spettro) come la trasformata di Fourier

$$(7.17) \quad S(t, \omega) = \int_{-\infty}^{\infty} R(t, \tau) \cdot e^{-j\omega\tau} d\tau$$

Sebbene lo spettro di potenza contenga esattamente la stessa informazione dell'autocorrelazione, spesso quest'ultimo è più facilmente interpretabile.

Lo spettro di potenza, più correttamente la **densità spettrale di potenza**, indica il contenuto "in potenza" (cioè quadratico) di un segnale. Ne parleremo estesamente in un prossimo capitolo. Lì discuteremo anche il senso delle frequenze negative.

Il caso di due processi

Supponiamo di avere due processi $X(t)$ e $Y(t)$, che per generalità consideriamo complessi.

Possiamo definire la **correlazione incrociata**¹⁸ (**cross-correlation**)

$$(7.18) \quad R_{xy}(t_1, t_2) = E[X(t_1) \cdot Y^*(t_2)] = R_{yx}^*(t_1, t_2)$$

e la **covarianza incrociata** (**cross-covariance**)

$$(7.19) \quad C_{xy}(t_1, t_2) = R_{xy}(t_1, t_2) - \mu_x(t_1) \cdot \mu_y(t_2)$$

Diciamo che i due processi sono **scorrelati** se

$$(7.20) \quad R_{xy}(t_1, t_2) = \mu_x(t_1) \cdot \mu_y(t_2)$$

e diciamo che sono **ortogonali** se

$$(7.21) \quad R_{xy}(t_1, t_2) = 0$$

¹⁸ Detta anche correlazione mutua, intercorrelazione o cross-correlazione.

A partire dalla correlazione incrociata $R_{xy}(t_1, t_2) = R_{xy}(t_1, t_1 + \tau) = R_{xy}(t, t + \tau)$, possiamo definire lo **spettro incrociato (cross-spectrum)** come

$$(7.22) \quad S_{xy}(t, \omega) = \int_{-\infty}^{\infty} R_{xy}(t, \tau) \cdot e^{-j\omega\tau} d\tau = S_{yx}^*(t, \omega)$$

Stazionarietà, ergodicità

Diciamo che un processo $X(t)$ è stazionario¹⁹ se le funzioni che lo definiscono sono indipendenti da uno spostamento nel tempo di qualsiasi ampiezza. Cioè se per qualsiasi T $X(t)$ è statisticamente indistinguibile da $X(t+T)$. Si ha cioè

$$(7.23) \quad F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = F(x_1, x_2, \dots, x_n; t_1 + T, t_2 + T, \dots, t_n + T)$$

e

$$(7.24) \quad f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f(x_1, x_2, \dots, x_n; t_1 + T, t_2 + T, \dots, t_n + T)$$

Ne deduciamo che per un processo stocastico stazionario si ha

$$f(x; t) = f(x)$$

(7.25)

$$\mu(t) = \mu = \text{costante}$$

$\mu = E[x(t)]$ è anche detto **cumulante del primo ordine**.

ponendo $\tau = t_2 - t_1$, abbiamo

$$f(x_1, x_2; t_1, t_2) = f(x_1, x_2; \tau)$$

(7.26)

$$R(t_1, t_2) = R(\tau) = R^*(-\tau)$$

$$C(t_1, t_2) = C(\tau) = C^*(-\tau)$$

¹⁹ Questa è la definizione di un processo stocastico stazionario “in senso stretto”. Si definisce anche la stazionarietà “in senso lato”.

$R_{xx}(\tau) = R(\tau) = E[x^*(t) \cdot x(t+\tau)]$ è, nell'ipotesi che sia $E[x(t)] = 0$, anche il **cumulante del secondo ordine** $C_{xx}(\tau)$.

Notiamo che

$$(7.27) \quad R_{xx}(0) - \mu^2 = \sigma^2$$

è la varianza del processo.

Lo spettro di potenza in questo caso perde la dipendenza dal tempo e viene definito come

$$(7.28) \quad S(\omega) = \int_{-\infty}^{\infty} R(\tau) \cdot e^{-j\omega\tau} d\tau$$

Possiamo definire la potenza totale del processo l'integrale totale su tutta la banda di frequenza (non pulsazione), ottenendo

$$(7.29) \quad R(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) d\omega = \int_{-\infty}^{\infty} S(2\pi \cdot \nu) d\nu$$

Possono definirsi cumulanti di altri ordini, che, nell'ipotesi che sia $E[x(t)] = 0$, sono nel caso del del terzo ordine

$$(7.30) \quad C_{xxx}(\tau_1, \tau_2) = E[x^*(t) \cdot x(t+\tau_1) \cdot x(t+\tau_2)]$$

e del quarto ordine

$$(7.31) \quad \begin{aligned} C_{xxxx}(\tau_1, \tau_2, \tau_3) = & E[x^*(t) \cdot x(t+\tau_1) \cdot x(t+\tau_2) \cdot x^*(t+\tau_3)] - \\ & - C_{xx}(\tau_1) \cdot C_{xx}(\tau_2 - \tau_3) - C_{xx}(\tau_2) \cdot C_{xx}(\tau_1 - \tau_3) - M_{xx}^*(\tau_3) \cdot M_{xx}(\tau_1 - \tau_2) \end{aligned}$$

dove

$$(7.32) \quad M_{xx}(\tau) = E[x(t) \cdot x(t+\tau)]$$

è il momento del secondo ordine (eguale al cumulante del secondo ordine nel caso di processi reali).

e notiamo che da $C_{xxx}(0,0)$ si ricava il momento del terzo ordine e da

$C_{xxxx}(0,0,0)$ quello del quart'ordine.

Tutti i cumulanti di processi stazionari reali sono simmetrici nei loro argomenti, cioè

(7.33)

$$C_{xx}(\tau) = C_{xx}(-\tau)$$

$$C_{xxx}(\tau_1, \tau_2) = C_{xxx}(\tau_2, \tau_1) = C_{xxx}(-\tau_1, \tau_2 - \tau_1)$$

$$C_{xxxx}(\tau_1, \tau_2, \tau_3) = C_{xxxx}(\tau_2, \tau_1, \tau_3) = C_{xxxx}(\tau_1, \tau_3, \tau_2) = C_{xxxx}(-\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_1)$$

La trasformata di Fourier di $C_{xxx}(\tau_1, \tau_2)$ è detto **bispettro** e si indica con $S_{xxx}(\omega_1, \omega_2)$,

$$(7.34) \quad S_{xxx}(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{xxx}(\tau_1, \tau_2) \cdot e^{-j\omega_1\tau_1} \cdot e^{-j\omega_2\tau_2} \cdot d\tau_1 \cdot d\tau_2$$

la trasformata di Fourier di $C_{xxxx}(\tau_1, \tau_2, \tau_3)$ è detto **trispettro** e si indica con $S_{xxxx}(\omega_1, \omega_2, \omega_3)$

$$(7.35) \quad S_{xxxx}(\omega_1, \omega_2, \omega_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{xxxx}(\tau_1, \tau_2, \tau_3) \cdot e^{-j\omega_1\tau_1} \cdot e^{-j\omega_2\tau_2} \cdot e^{-j\omega_3\tau_3} \cdot d\tau_1 \cdot d\tau_2 \cdot d\tau_3$$

Per due processi **congiuntamente stazionari (jointly stationary)** abbiamo

$$(7.36) \quad R_{xy}(t_1, t_2) = R_{xy}(\tau) = E[X(t+\tau) \cdot Y^*(t)]$$

Diciamo **ergodico** un processo in cui le statistiche di insieme sono eguali alle statistiche temporali. In altre parole possiamo conoscere tutte le proprietà del processo stocastico osservando la realizzazione effettiva. È intuitivo che un processo non può essere ergodico se non è stazionario.

La stazionarietà e l'ergodicità possono essere ristrette ad alcune caratteristiche del processo (per esempio la media, la varianza o l'autocorrelazione).

Esempi

Facciamo tre esempi (di processi continui):

1. Consideriamo un processo che contiene 3 funzioni del tempo

$$x_1(t) = 0 \quad \rightarrow \text{Prob. } \frac{1}{2}$$

$$x_2(t) = \sin(2\pi \cdot 50 \cdot t) \quad \rightarrow \text{Prob. } \frac{1}{4}$$

$$x_3(t) = \sin(2\pi \cdot 100 \cdot t) \quad \rightarrow \text{Prob. } \frac{1}{4}$$

Questo processo stocastico potrebbe essere il modello di un dispositivo elettrico che con probabilità 0.5 è acceso (e ovviamente con la stessa probabilità è spento) e quando è acceso può generare con eguale probabilità una sinusoide a 100 Hz o una a 50.

2. Consideriamo un processo che è composto di infinite funzioni del tempo definite da

$$x_A(t) = A \cdot \sin(10 \cdot t)$$

dove A è distribuita uniformemente tra 0 e 2.

Questo processo stocastico potrebbe essere il modello di un oscillatore elettrico che genera una tensione sinusoidale, con una manopola che ne stabilisce l'ampiezza, e la manopola può stare in una qualsiasi posizione con la stessa probabilità (ma non può essere cambiata di posizione).

3. Rumore Johnson - Consideriamo tutte le possibili funzioni del tempo che possono aversi come uscita di un amplificatore rumoroso chiuso su una certa resistenza ad una certa temperatura. Quello che osserviamo è il rumore Johnson, o meglio una delle possibili infinite realizzazioni di un processo stocastico con statistiche gaussiane.
4. Processo di Poisson. Consideriamo tutte le possibili funzioni del tempo che possono aversi come uscita di un contatore di particelle che rivela per esempio raggi cosmici. Anche in questo caso ci troviamo di fronte a un processo stocastico, anche se è più difficile conoscere la distribuzione di probabilità delle varie realizzazioni (e in parecchi casi non serve, essendo sufficienti altre funzioni, più semplici, legate alla distribuzione di probabilità).

Processi normali

Chiamiamo normale (o gaussiano) un processo stocastico. in cui sono normali le funzioni di densità di qualsiasi ordine.

Ci sono due motivi per cui i processi stocastici normali sono di grande importanza:

1. Molto spesso fenomeni naturali (e non) sono ben rappresentati da processi stocastici normali
2. Grazie alle proprietà matematiche che appresso enunceremo, si può costruire una semplice ed elegante teoria dei processi stocastici normali, per cui, anche se i fenomeni di interesse non sono esattamente rappresentabili con processi normali, si preferisce approssimarli, quando possibile, con essi.

Ecco le due proprietà fondamentali:

PROPRIETÀ I - Per un processo stocastico normale la funzione densità di ordine n (per qualsiasi n) è determinata se sono note la media $\mu(t)$ e l'autocorrelazione $R(t_1, t_2)$.

Come sappiamo, una trasformazione lineare di una variabile casuale normale produce una variabile casuale normale. Analogamente abbiamo che:

PROPRIETÀ II - Una trasformazione lineare di un processo stocastico normale produce un processo stocastico normale.

Cioè se facciamo passare un processo normale attraverso un sistema lineare, otteniamo in uscita un processo normale.

Particolarmente semplice e illuminante è lo sviluppo nell'ipotesi stazionaria, che d'ora in poi sottintenderemo verificata.

Il rumore bianco

Il rumore bianco. è un processo stocastico normale stazionario (ed ergodico) definito da

$$\mu(t) = 0$$

(7.37)

$$R(\tau) = \delta(\tau)$$

Cioè due valori del processo a due istanti qualsiasi, comunque vicini, sono scorrelati (e indipendenti: ricordiamo che due variabili casuali normali con correlazione nulla sono indipendenti).

È immediato vedere che la densità spettrale di potenza è

$$(7.38) \quad S(\omega) = \int_{-\infty}^{\infty} R(\tau) \cdot e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} \delta(\tau) \cdot e^{-j\omega\tau} d\tau = 1$$

Lo spettro quindi contiene tutte le frequenze, con la stessa ampiezza. Comprendiamo allora il motivo del nome “bianco”: è in analogia con la luce bianca, che è composta da tutti i colori, cioè tutte le frequenze (in realtà solo del visibile e non tutte con la stessa potenza)²⁰.

Notiamo che il rumore bianco, a parte altre caratteristiche “ideali”, è caratterizzato da potenza infinita. Ciò per la presenza di potenza a tutte le frequenze. Tutti i sistemi fisici sono limitati in frequenza, e quindi tutti i segnali fisici, che sono da essi generati, lo sono.

²⁰ Talvolta si parla di rumore “rosa” per indicare un rumore con spettro abbastanza piatto alle basse frequenze e praticamente nullo alle alte. Questo, a differenza del rumore bianco, è un rumore “possibile”.

Trasformazioni di processi stocastici

Dato un processo stocastico X , possiamo assegnare secondo certe regole a ciascuna funzione del tempo $x(t)$ del processo, una funzione $y(t)$, ottenendo un nuovo processo stocastico Y .

Y può essere considerato l'uscita di un opportuno sistema.

Far passare un processo stocastico X attraverso un sistema F significa costruire un nuovo processo Y in cui tutte le funzioni del tempo $x(t)$ sono sostituite dalle uscite $y(t)$ di F alle $x(t)$.

Consideriamo alcuni casi particolari.

Sistema statico (senza memoria)

Sia in tal caso

$$(7.39) \quad Y(t) = g(X(t))$$

Le funzioni di densità del processo Y sono ricavabili con le tecniche di trasformazione delle variabili aleatorie. Si ha

$$(7.40) \quad f_y(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n) = \frac{f_x(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{|g'(x_1)| \cdot |g'(x_2)| \cdot \dots \cdot |g'(x_n)|}$$

con $y_i = g(x_i)$.

Per quanto riguarda il valor medio e l'autocorrelazione abbiamo

$$(7.41) \quad E[Y(t)] = \int_{-\infty}^{\infty} g(x) \cdot f_x(x;t) \cdot dx$$

e

$$(7.42) \quad R_{yy}(t_1, t_2) = E[y(t_1) \cdot y(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1) \cdot g(x_2) \cdot f_x(x_1, x_2; t_1, t_2) \cdot dx_1 \cdot dx_2$$

Inoltre, se X è stazionario, anche Y lo è.

Sistema lineare (tempo invariante)

Indichiamo con L l'operatore lineare che descrive il sistema. (vedi (6.4)). Si hanno i seguenti risultati:

$$(7.43) \quad E[Y(t)] = L[E[X(t)]]$$

$$(7.44) \quad R_{xy}(t_1, t_2) = L_{t_2}[R_{xx}(t_1, t_2)]$$

$$(7.45) \quad R_{yy}(t_1, t_2) = L_{t_1}[R_{xy}(t_1, t_2)]$$

Con L_{t_2} indichiamo "rispetto alla variabile t_2 ", considerandola variabile corrente e t_1 parametro; analogamente per L_{t_1} .

Se il processo in ingresso a un sistema lineare tempo invariante è stazionario, lo è anche quello in uscita.

Nell'ipotesi di stazionarietà, supponiamo di avere un sistema (continuo) definito da una funzione di trasferimento $F(s)$, ovvero dalla risposta impulsiva $f(t)$. Per il valor medio si ha

$$(7.46) \quad E[Y(t)] = \mu_Y = E\left[\int_{-\infty}^{\infty} X(t-\theta) f(\theta) \cdot d\theta\right] = \mu_X \cdot \int_{-\infty}^{\infty} f(\theta) \cdot d\theta = \mu_X \cdot F(0)$$

Per calcolare l'autocorrelazione dell'uscita è opportuno calcolare prima la correlazione incrociata tra l'uscita e l'ingresso:

$$(7.47) \quad \begin{aligned} R_{YX}(\tau) &= E[Y(t) \cdot X^*(t-\tau)] = E\left[\int_{-\infty}^{\infty} x(t-\theta) \cdot x^*(t-\tau) \cdot f(\theta) d\theta\right] = \\ &= \int_{-\infty}^{\infty} R_{XX}(\tau-\theta) \cdot f(\theta) d\theta \end{aligned}$$

e

$$(7.48) \quad R_{XY}(\tau) = \int_{-\infty}^{\infty} R_{XX}(\tau-\theta) \cdot f^*(-\theta) d\theta$$

e quindi

$$(7.49) \quad \begin{aligned} R_{YY}(\tau) &= E[Y(t+\tau) \cdot Y^*(t)] = E\left[\int_{-\infty}^{\infty} y(t+\tau) \cdot x^*(t-\theta) \cdot f^*(\theta) d\theta\right] = \\ &= \int_{-\infty}^{\infty} R_{YX}(\tau+\theta) \cdot f^*(\theta) d\theta \end{aligned}$$

Ora, ricordando la proprietà della convoluzione e la definizione di spettro di potenza, troviamo

$$(7.50) \quad S_{XY}(\omega) = S_{XX}(\omega) \cdot F^*(j\omega)$$

e

$$(7.51) \quad S_{yy}(\omega) = S_{xx}(\omega) \cdot |F(j\omega)|^2$$

Lo stesso risultato per gli spettri possiamo ottenerlo ragionando sul significato di spettro di potenza e di funzione di trasferimento.

Infatti, in una banda molto piccola intorno a ω_0 il segnale è approssimabile con una sinusoide con una certa ampiezza A e una certa fase; lo spettro di potenza è kA^2 , dove k è una opportuna costante, e indipendente dalla fase. Il passaggio attraverso il sistema lineare modifica ovviamente tale valore di un fattore $|F(j\omega_0)|^2$, come descritto dalla (7.51). Analogamente si può ragionare per la (7.50).

Un caso particolare: il derivatore

Il **derivatore**, che esegue la derivata di un processo stocastico, spesso detta **derivata stocastica** (ovviamente le funzioni che compongono il processo devono essere derivabili). Abbiamo in questo caso

$$(7.52) \quad E[x'(t)] = \frac{dE[x(t)]}{dt}$$

$$(7.53) \quad R_{xx'}(t_1, t_2) = \frac{\partial R_{xx}(t_1, t_2)}{\partial t_2}$$

$$(7.54) \quad R_{x'x'}(t_1, t_2) = \frac{\partial R_{xx'}(t_1, t_2)}{\partial t_1} = \frac{\partial^2 R_{xx}(t_1, t_2)}{\partial t_1 \partial t_2}$$

e se il processo è stazionario, con $\tau = t_1 - t_2$,

$$(7.55) \quad R_{xx'}(\tau) = -\frac{dR_{xx}(\tau)}{d\tau}$$

$$(7.56) \quad R_{x'x'}(\tau) = \frac{dR_{xx'}(\tau)}{d\tau^2} = -\frac{d^2R_{xx}(\tau)}{d\tau^2}$$

Processi stocastici normali

Proprietà fondamentali

Chiamiamo **normale** (o **gaussiano**) un processo stocastico. in cui sono normali le funzioni di densità di qualsiasi ordine.

Ci sono due motivi per cui i processi stocastici normali sono di grande importanza:

1. Molto spesso fenomeni naturali (e non) sono ben rappresentati da processi stocastici normali
2. Grazie alle proprietà matematiche che appresso enunceremo, si può costruire una semplice ed elegante teoria dei processi stocastici normali, per cui, anche se i fenomeni di interesse non sono esattamente rappresentabili con processi normali, si preferisce approssimarli, quando possibile, con essi.

Ecco le due proprietà fondamentali:

PROPRIETÀ I - Per un processo stocastico normale la funzione densità di ordine n (per qualsiasi n) è determinata se sono note la media $\mu(t)$ e l'autocorrelazione $R(t_1, t_2)$ (o l'autocovarianza $C(t_1, t_2) = R(t_1, t_2) - \mu(t_1) \cdot \mu(t_2)$).

Infatti, definita la matrice \mathbf{M} di covarianza tra le n $x(t_i)$ con elementi

$$(7.57) \quad m_{ij} = C(t_i, t_j)$$

abbiamo

$$(7.58) \quad f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{M}|}} \cdot \exp \left[-\frac{(\mathbf{x} - \boldsymbol{\mu}) \cdot \mathbf{M}^{-1} \cdot (\mathbf{x}^T - \boldsymbol{\mu}^T)}{2} \right]$$

dove $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ e $\boldsymbol{\mu} = \{\mu(t_1), \mu(t_2), \dots, \mu(t_n)\}$.

Come sappiamo, una trasformazione lineare di una variabile casuale normale produce una variabile casuale normale. Analogamente abbiamo che:

PROPRIETÀ II - Una trasformazione lineare di un processo stocastico normale produce un processo stocastico normale.

Cioè se facciamo passare un processo normale attraverso un sistema lineare, otteniamo in uscita un processo normale.

Particolarmente semplice e illuminante è lo sviluppo nell'ipotesi stazionaria, che d'ora in poi sottintenderemo verificata.

Processi stocastici discreti

Possiamo definire il concetto di processo stocastico anche nel caso in cui il tempo è discretizzato. Le proprietà sono analoghe. Qui considereremo solo il caso dei processi normali stazionari (e quindi anche ergodici).

Analogamente al caso continuo, possiamo definire lo spettro di potenza come la trasformata di Fourier per dati discreti (vedi (4.50) dell'autocorrelazione $R_{xx}(i)$)

$$(7.59) \quad S_x(\Omega) = \sum_{i=-\infty}^{\infty} R_{xx}(i) \cdot e^{-j \cdot i \cdot \Omega}$$

e lo spettro incrociato dalla correlazione incrociata $R_{xy}(i)$

$$(7.60) \quad S_{xy}(\Omega) = \sum_{i=-\infty}^{\infty} R_{xy}(i) \cdot e^{-j \cdot i \cdot \Omega}$$

Ovviamente vale la formula di inversione per la trasformata di Fourier.

Se invece della pulsazione normalizzata Ω vogliamo usare quella fisica $\omega = \frac{\Omega}{\Delta t}$, la densità spettrale va opportunamente normalizzata

$$(7.61) \quad S'(\omega) = S(\Omega) \cdot \Delta t$$

Ciò garantisce la proprietà fondamentale che l'integrale sull'intera banda della pulsazione ($\frac{2\pi}{\Delta t}$ nel caso di ω , 2π nel caso di Ω) sia pari alla varianza del segnale moltiplicato 2π . Se l'integrale viene fatto nella frequenza invece che nella pulsazione, si ha semplicemente la varianza del segnale.

Rumore bianco discreto

Anche nel caso discreto possiamo definire il rumore gaussiano bianco. Esso consiste in una successione di campioni distribuiti normalmente con valor medio nullo (e spesso si prende con varianza unitaria).

L'autocorrelazione è

$$(7.62) \quad R(j) = \sigma^2 \cdot \delta_j$$

e lo spettro, usando la (4.50), è

$$(7.63) \quad S(\Omega) = \sigma^2$$

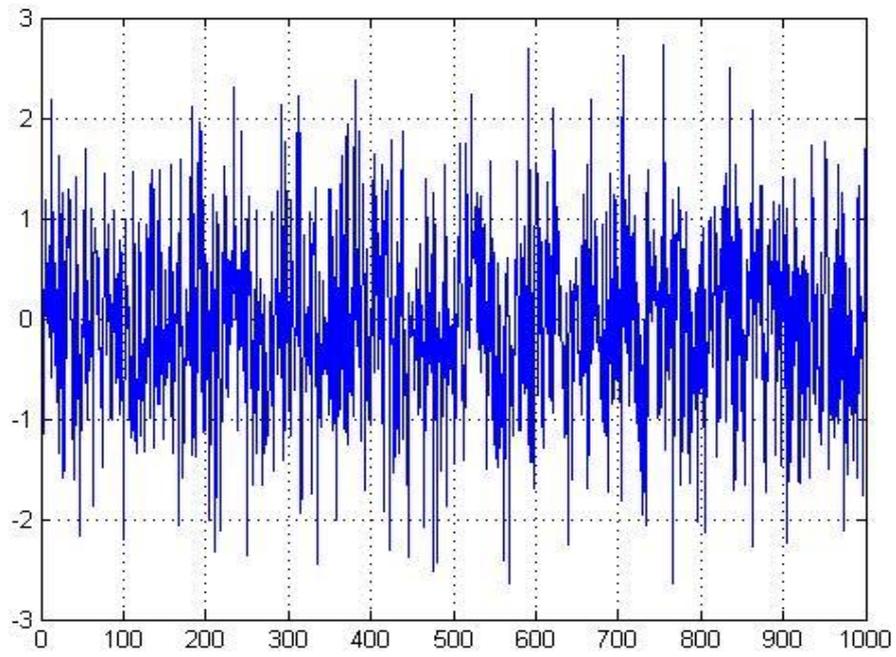
Se vogliamo “lavorare” con le frequenze fisiche, allora

$$(7.64) \quad S'(\omega) = S(\Omega) \cdot \Delta t = \sigma^2 \cdot \Delta t$$

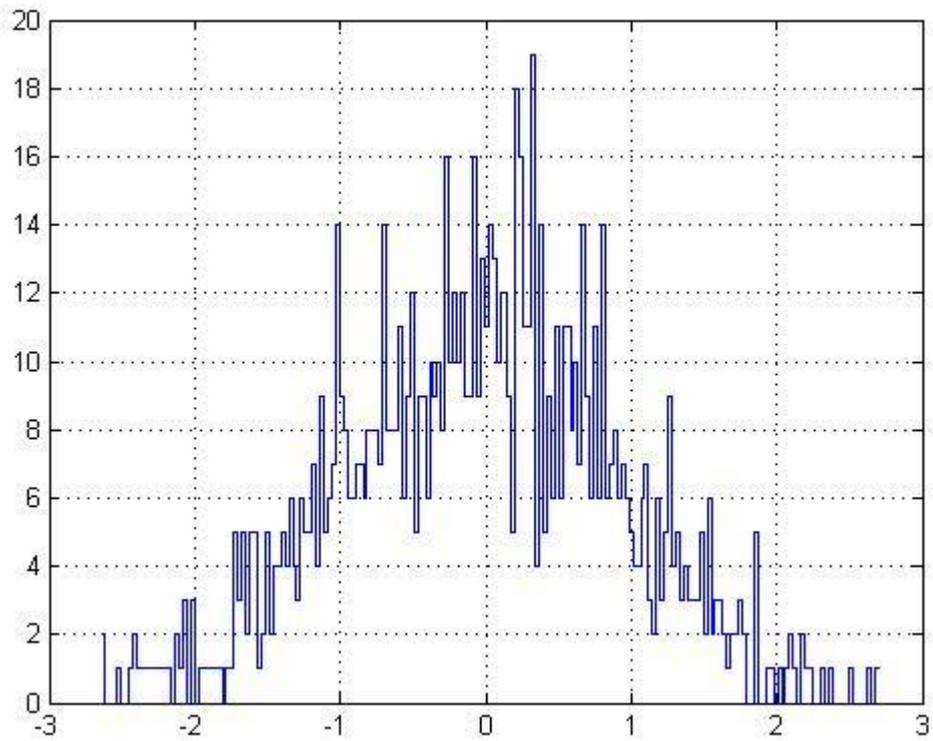
dove $0 \leq \Omega < 2\pi$, ovvero $-\pi \leq \Omega < \pi$. Come si vede, lo spettro è piatto come nel caso del rumore bianco continuo, ma la banda del segnale è limitata, quindi, a differenza del rumore bianco continuo, questo non crea paradossi, è fisicamente realizzabile: infatti non è altro che una successione di valori distribuiti gaussianamente con varianza σ^2 , indipendenti.

Si noti come la densità spettrale “fisica” si abbassa proporzionalmente al tempo di campionamento.

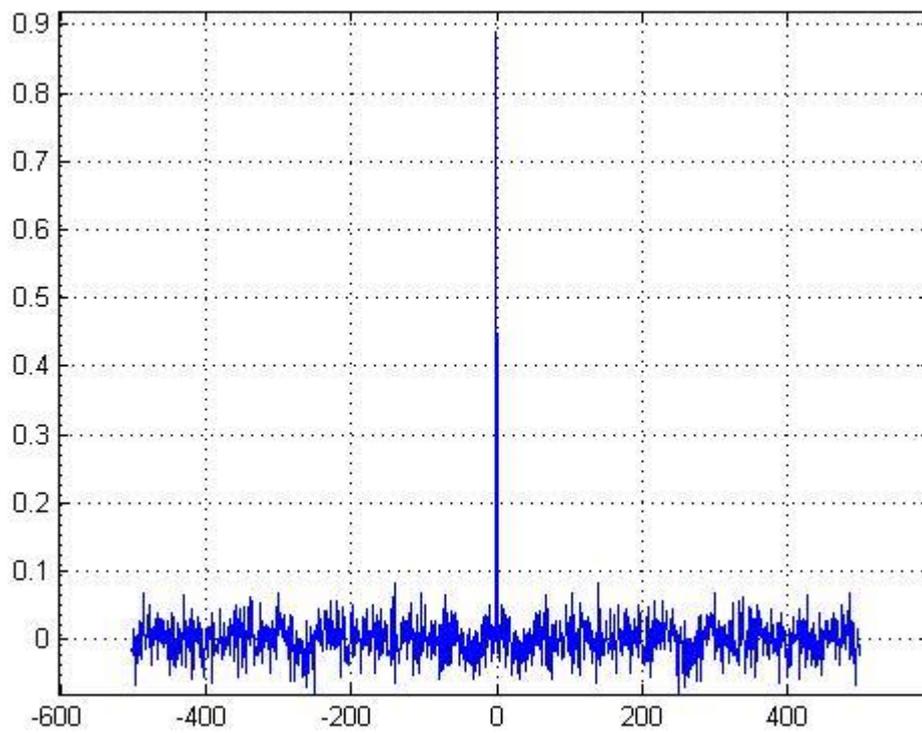
Ecco una realizzazione di rumore bianco discreto (1000 campioni)



ed eccone l'istogramma



e la (stima della) autocorrelazione



Processi normali e sistemi discreti

Supponiamo di avere un sistema lineare definito da una risposta impulsiva w_i e mandiamo in ingresso un rumore bianco n_i :

$$(7.65) \quad y_i = \sum_j n_{i-j} \cdot w_j$$

L'autocorrelazione è data da

$$(7.66) \quad R_{yy}(k) = E[y_i y_{i+k}^*] = E\left[\left(\sum_j n_{i-j} \cdot w_j\right)\left(\sum_j n_{i+k-j} \cdot w_j^*\right)\right]$$

Poiché l'autocorrelazione del rumore è una delta, abbiamo

$$(7.67) \quad R_{yy}(k) = \sigma_n^2 \sum_j w_j w_{j+k}^*$$

Quindi, a parte un coefficiente dato dalla varianza del rumore bianco, l'autocorrelazione è ottenuta dalla convoluzione della risposta impulsiva con se stessa invertita temporalmente.

Quindi, se è $F(z)$ la funzione di trasferimento del sistema, lo spettro di potenza, cioè la trasformata di Fourier dell'autocorrelazione, è

$$(7.68) \quad S(\Omega) = \sigma_n^2 \cdot |F(z)|^2 \Big|_{z=\exp(j\Omega)}$$

Résumé dei risultati

La teoria dei processi stocastici è stata fin qui sviluppata p.er lo più per i processi continui. Analoghi sviluppi possono farsi nel caso dei processi discreti. In particolare per quanto riguarda la teoria dei processi normali e i sistemi lineari, per i processi discreti si ottengono analoghi risultati considerando sistemi discreti. Sintetizziamo qui i risultati più imporanti.

Analogamente al caso continuo, possiamo definire lo spettro di potenza come la trasformata di Fourier per dati discreti (vedi (4.50) dell'autocorrelazione $R_{xx}(i)$)

$$(7.69) \quad S_x(\Omega) = \sum_{i=-\infty}^{\infty} R_{xx}(i) \cdot e^{-j \cdot i \cdot \Omega}$$

e lo spettro incrociato dalla correlazione incrociata $R_{xy}(i)$

$$(7.70) \quad S_{xy}(\Omega) = \sum_{i=-\infty}^{\infty} R_{xy}(i) \cdot e^{-j \cdot i \cdot \Omega}$$

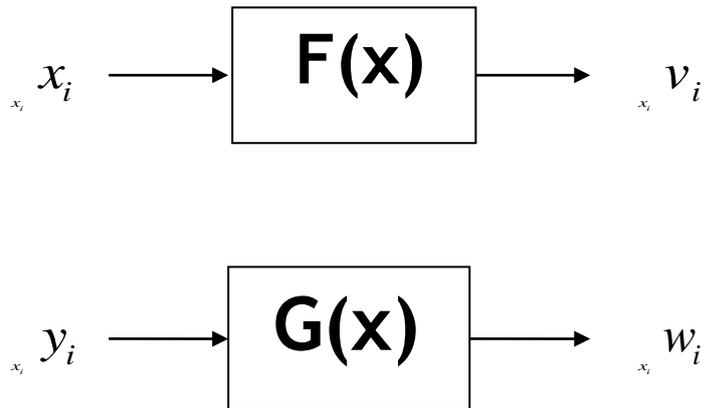
Ovviamente vale la formula di inversione per la trasformata di Fourier.

Se invece della pulsazione normalizzata Ω vogliamo usare quella fisica $\omega = \frac{\Omega}{\Delta t}$, la densità spettrale va opportunamente normalizzata

$$(7.71) \quad S'(\omega) = S(\Omega) \cdot \Delta t$$

Ciò garantisce la proprietà fondamentale che l'integrale sull'intera banda della pulsazione ($\frac{2\pi}{\Delta t}$ nel caso di ω , 2π nel caso di Ω) sia pari alla varianza del segnale moltiplicato 2π . Se l'integrale viene fatto nella frequenza invece che nella pulsazione, si ha semplicemente la varianza del segnale.

Consideriamo lo schema seguente:



In esso x_i, y_i, v_i, w_i sono processi stocastici normali discreti.

Dalla (6.27) abbiamo

$$(7.72) \quad F(z) = \sum_{i=0}^{\infty} f_i \cdot z^{-i}$$

$$G(z) = \sum_{i=0}^{\infty} g_i \cdot z^{-i}$$

e quindi

$$(7.73) \quad v_i = \sum_{k=0}^{\infty} f_k \cdot x_{i-k}$$

$$w_i = \sum_{k=0}^{\infty} g_k \cdot y_{i-k}$$

Allora

$$(7.74) \quad R_{vx}(j) = E[v_{i+j} \cdot x_i^*] = E\left[\sum_{k=0}^{\infty} f_k \cdot x_{i-k+j} \cdot x_i^*\right] = \sum_{k=0}^{\infty} f_k \cdot R_{xx}(j)$$

$$(7.75) \quad R_{vy}(j) = E[v_{i+j} \cdot y_i^*] = E\left[\sum_{k=0}^{\infty} f_k \cdot x_{i-k+j} \cdot y_i^*\right] = \sum_{k=0}^{\infty} f_k \cdot R_{xy}(j)$$

e poiché $R_{xy}(j) = R_{yx}^*(-j)$,

.....

e inoltre

$$(7.76) \quad R_{vy}(j) = E[v_{i+j} \cdot y_i^*] = E\left[\sum_{k=0}^{\infty} f_k \cdot x_{i-k+j} \cdot y_i^*\right] = \sum_{k=0}^{\infty} f_k \cdot R_{xy}(j)$$

e

$$(7.77) \quad R_{vw}(j) =$$

Per passare agli spettri si usa la proprietà della convoluzione e si ha

$$(7.78) \quad S_{vy}(z) = F(z) \cdot S_{xy}(z)$$

$$(7.79) \quad S_{vw}(z) = G(z^{-1}) \cdot S_{vy}(z) = F(z) \cdot G(z^{-1}) \cdot S_{xy}(z)$$

che comprende come caso particolare (per $x_i = y_i$ e $F=G$)

$$(7.80) \quad S_{vx}(z) = F(z) \cdot S_{xx}(z)$$

$$(7.81) \quad S_{vv}(z) = F(z^{-1}) \cdot S_{xx}(z) = F(z) \cdot F(z^{-1}) \cdot S_{xx}(z) = |F(z)|^2 \cdot S_{xx}(z)$$

Processi AR, MA e ARMA

Se mettiamo in ingresso a un sistema lineare discreto, di funzione di trasferimento $F(z)$, del rumore bianco di varianza σ^2 , produciamo un processo stocastico normale con media nulla e spettro di potenza

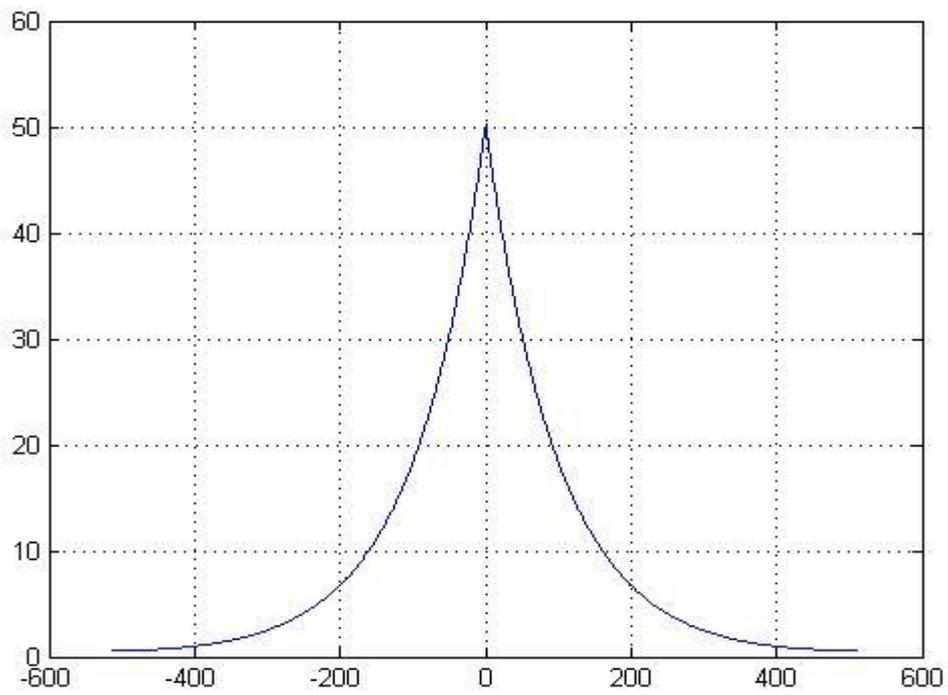
$$(7.82) \quad S(\Omega) = \sigma^2 \cdot |F(e^{j\Omega})|^2$$

Se il sistema è un sistema AR o MA o ARMA, diremo che il processo è rispettivamente AR, MA o ARMA. Mostriamo ora alcuni casi.

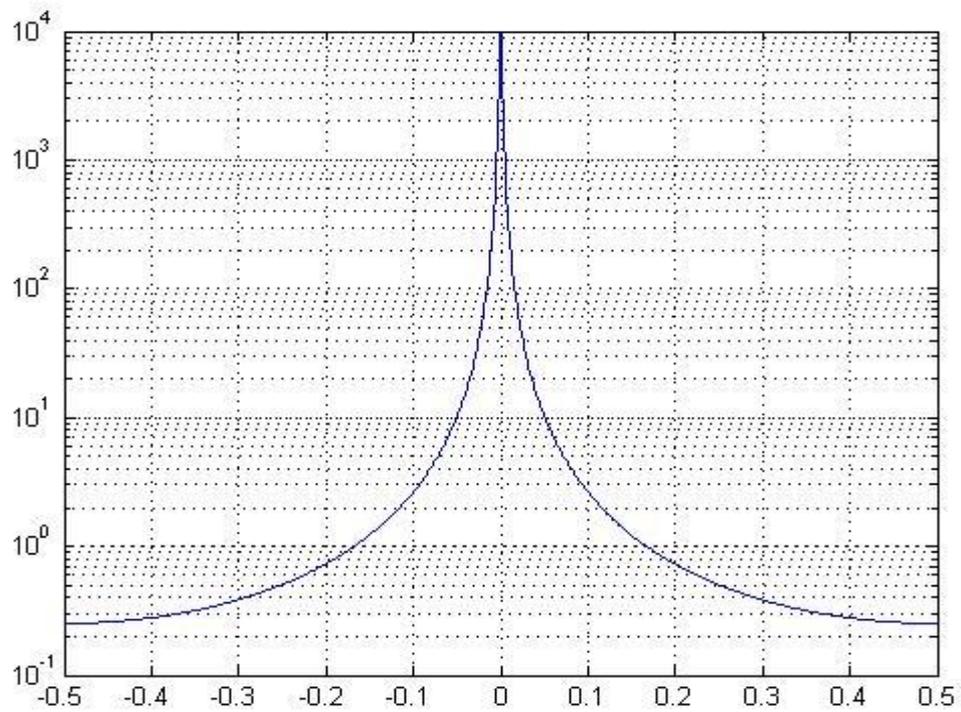
Processo AR del primo ordine (reale)

Consideriamo il caso in cui $b_0=1$, $a_0=1$, $a_1=-0.99$, essendo 1 il tempo di campionamento.. Ecco i grafici dell'autocorrelazione, dello spettro e 1000 campioni di una realizzazione

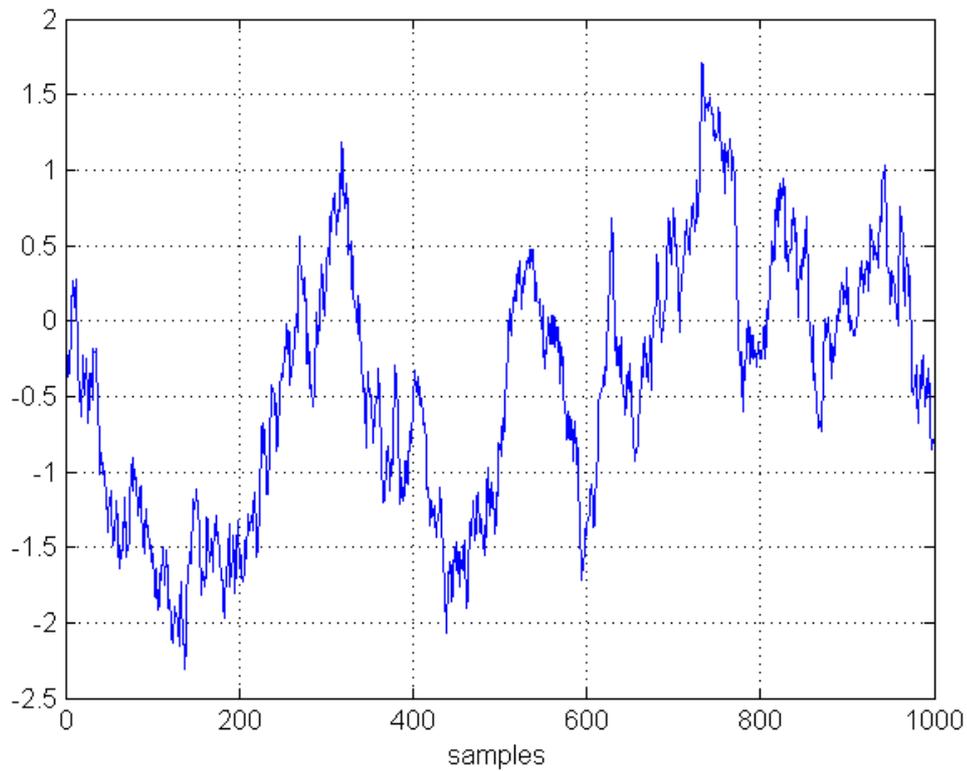
Autocorrelazione



Spettro (in ascissa la frequenza)



Un pezzo di una realizzazione



Processo AR del primo ordine (complesso)

Consideriamo il caso in cui

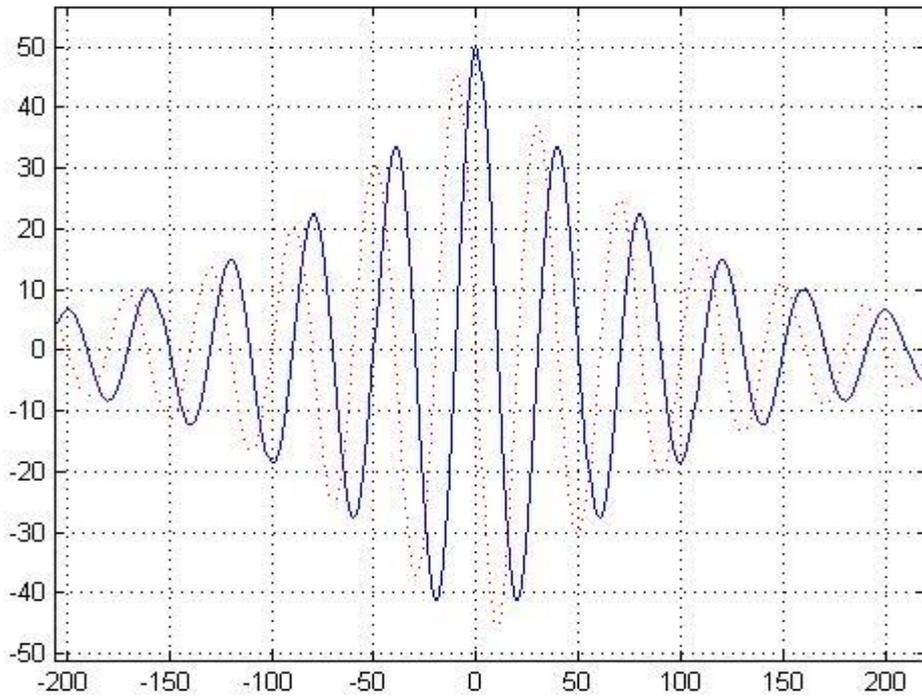
$$b_0 = 1, a_0 = 1, a_1 = 0.99 \cdot \exp(-j \cdot \pi / 20) = (-0.9778 + j \cdot 0.1549)$$

essendo 1 il tempo di campionamento. Ecco i grafici dell'autocorrelazione, dello spettro e 1000 campioni di una realizzazione.

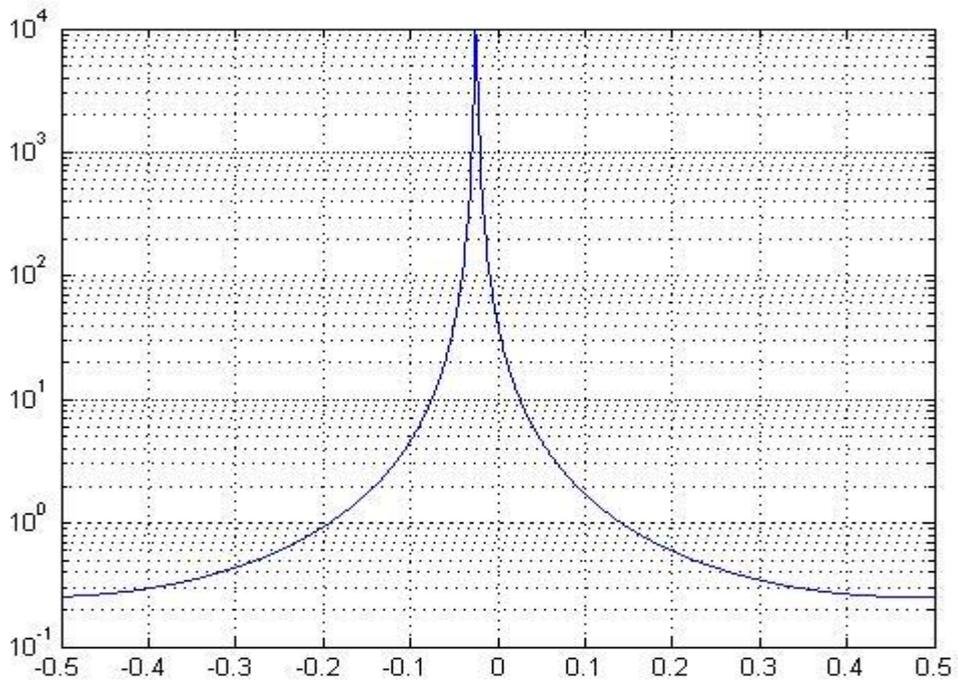
Si noti che in questo esempio il tempo di decadimento è lo stesso dell'esempio precedente, ma abbiamo una risonanza a frequenza negativa, a $\nu = -\frac{1}{40} = -0.025$.

Possiamo dire che questo processo è una rotazione oraria di $360/40=9$ gradi nel piano di Gauss del processo reale precedente.

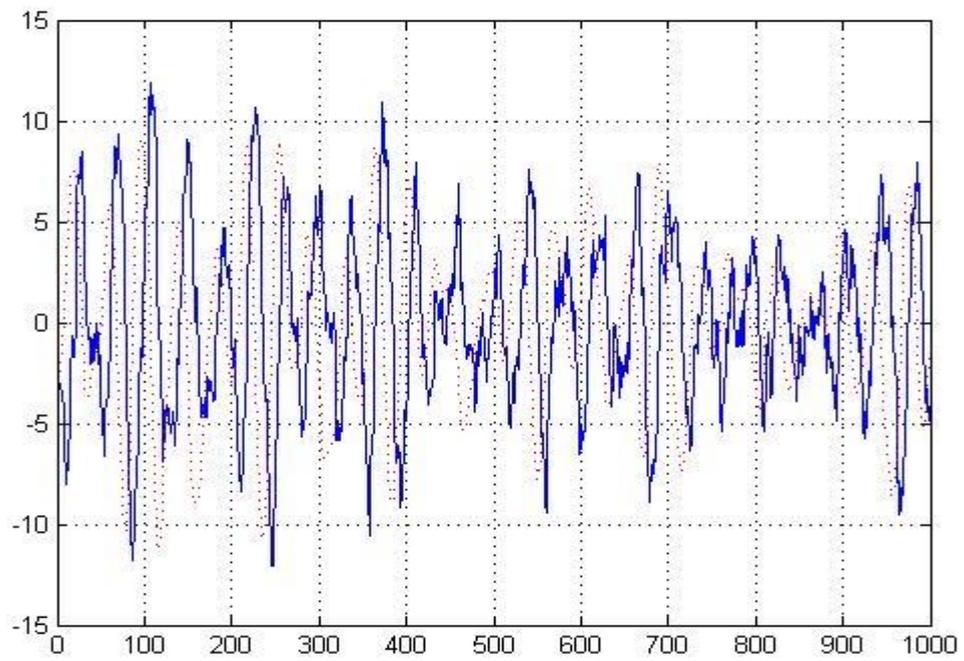
Autocorrelazione (la parte immaginaria è punteggiata)



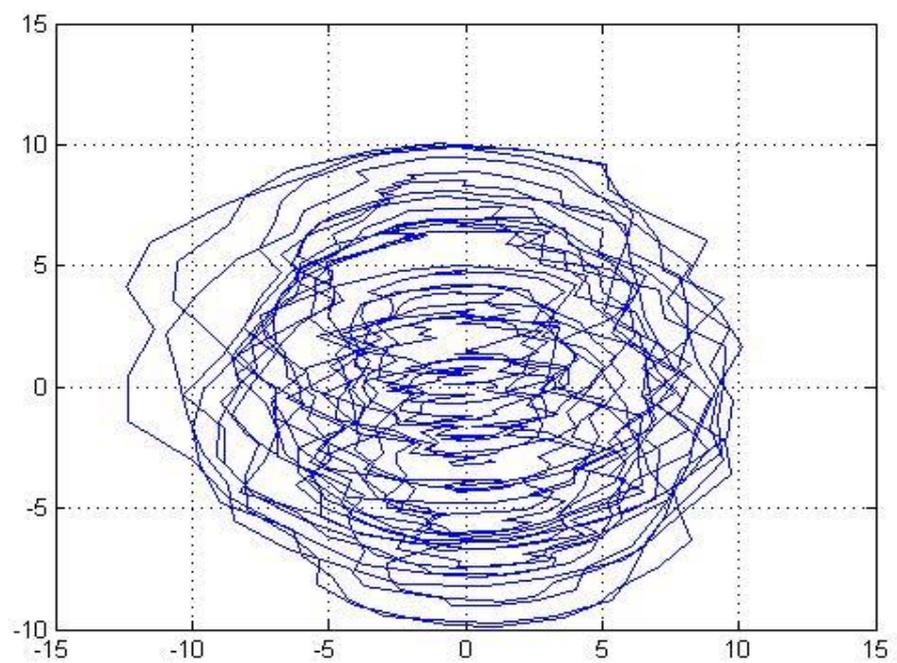
Spettro



Un pezzo di una realizzazione (la parte immaginaria è punteggiata)



Parte immaginaria vs reale



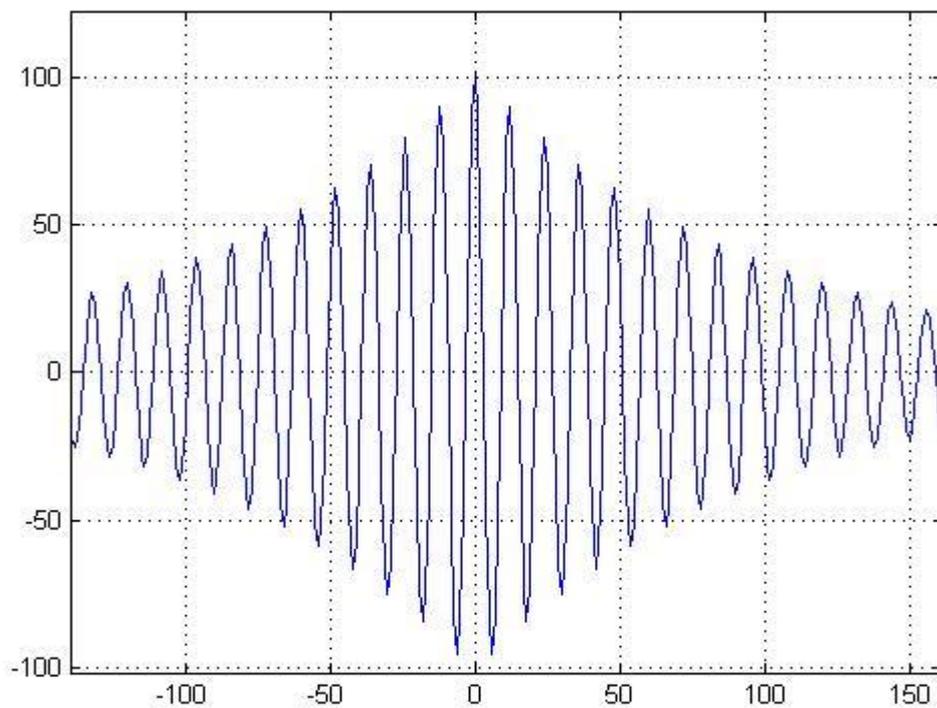
Processo AR del secondo ordine

Consideriamo il caso in cui

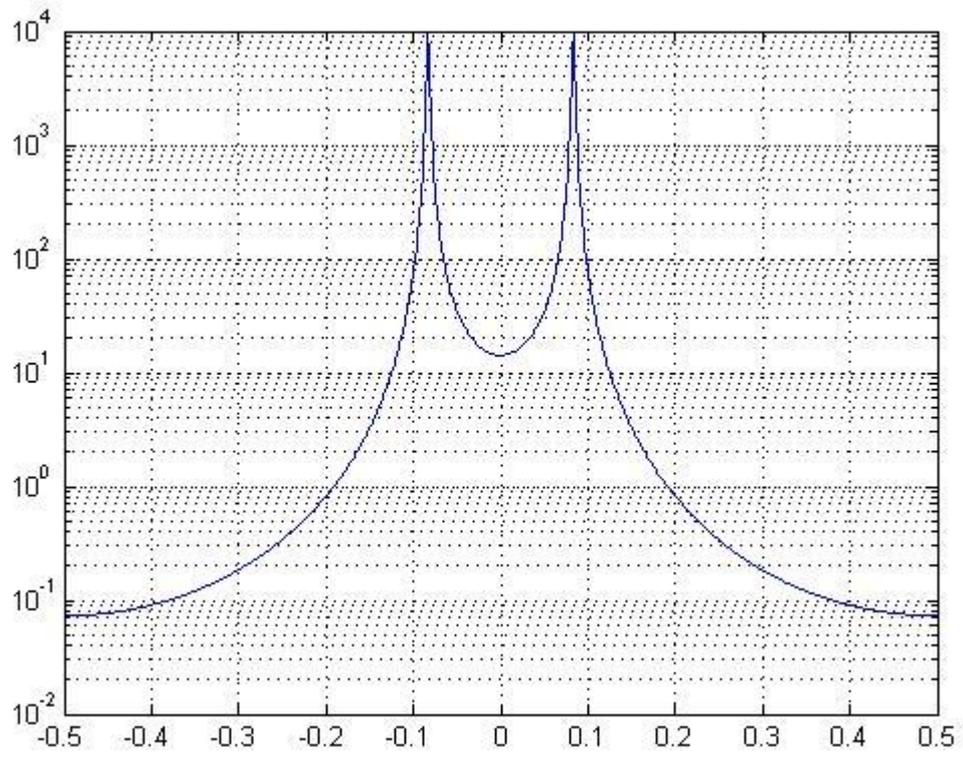
$$b_0 = 1, a_0 = 1, a_1 = -2 \cdot 0.99 \cdot \cos\left(\frac{\pi}{6}\right) = -1.7147, a_2 = r^2 = 0.9801$$

Abbiamo

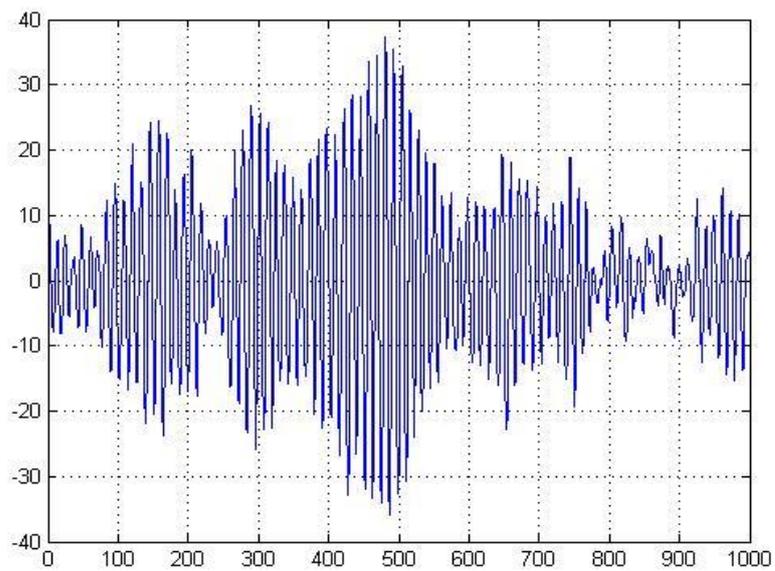
Autocorrelazione



Spettro



Pezzo di dati



Processo di Poisson

Analizziamo un altro processo stocastico, di grande interesse non solo in Fisica, il processo di Poisson, che, come suggerisce il nome, è collegato alla distribuzione di Poisson.

Il problema è quello di modellare un processo fisico che consiste nella generazione di eventi, tutti eguali, distribuiti in modo uniforme nel tempo.

Un esempio è il decadimento radioattivo di un campione di materiale; gli eventi sono i singoli eventi rivelati da un contatore. In questo caso il tempo di osservazione deve essere molto più breve del tempo di decadimento dei nuclei del materiale, altrimenti gli eventi non sono distribuiti uniformemente nel tempo.

Un altro esempio di interesse fisico è il passaggio di una giunzione da parte di elettroni o lacune. Questo processo, in cui il numero di eventi è molto grande nell'unità di tempo²¹, è la causa del rumore shot.

Sebbene si possa definire opportunamente anche nel discreto, il processo di Poisson è un processo stocastico continuo.

Le funzioni del tempo del processo possono essere la successione di delte di Dirac

$$x(t) = \sum_{i=-\infty}^{\infty} \delta(t-t_i)$$

dove con t_i si indicano gli istanti di occorrenza degli eventi. Spesso come funzioni del processo vengono prese le funzioni non decrescenti a valore intero che contano il numero di arrivi prima del tempo t : in pratica l'integrale della $x(t)$ di **Error! Reference source not found.**, da 0 a t .

Il processo ha un unico parametro, indicato con λ , che indica la densità degli eventi, cioè il numero atteso di eventi nell'unità di tempo. Nel periodo di tempo T si aspetta il numero di eventi

$$(7.83) \quad \mu = \lambda \cdot T$$

La distribuzione degli eventi in tale periodo è data dalla distribuzione di Poisson

²¹ Il passaggio della giunzione non è istantaneo, quindi, se il numero di eventi è elevato, dalla corrente risultante possono non essere "risolti" i singoli eventi.

$$(7.84) \quad P(k) = \frac{(\lambda \cdot T)^k}{k!} e^{-\lambda \cdot T}$$

e il numero di eventi in due intervalli disgiunti sono due variabili casuali indipendenti.

Si può calcolare che il tempo τ tra due eventi successivi ha densità di probabilità esponenziale

$$(7.85) \quad f(\tau) = \lambda \cdot u(\tau) \cdot e^{-\lambda \cdot \tau}$$

con valor medio

$$(7.86) \quad E[\tau] = \frac{1}{\lambda}$$

Non sempre il processo di Poisson con λ costante è un buon modello per processi fisici a eventi (in cui la cosa importante è il tempo di occorrenza degli eventi, non altre loro caratteristiche). Ciò perché la densità non è uniforme nel tempo, ma può variare; il processo cioè è descritto non più dal parametro λ costante, ma dalla funzione del tempo $\lambda(t)$. In tal caso possiamo calcolare il numero aspettato di eventi nel periodo $t_1 < t < t_2$ come

$$(7.87) \quad \mu = \int_{t_1}^{t_2} \lambda(t) \cdot dt$$

e la distribuzione è

$$(7.88) \quad P(k) = \frac{\left(\int_{t_1}^{t_2} \lambda(t) \cdot dt \right)^k}{k!} e^{-\int_{t_1}^{t_2} \lambda(t) \cdot dt}$$

8 - Analisi statistica dei segnali

Live Script

Caratteristiche statiche

Una stima della densità del primo ordine (vedi (7.2), che nel caso stazionario si riduce a $f(x)$) può ottenersi facendo un istogramma dei dati, normalizzato per il numero totale dei dati utilizzati.

Per fare questo istogramma occorre scegliere bene i bin, che non necessariamente saranno tutti uguali. Se si usano pochi bin, si hanno informazioni un po' "rozze" sulla $f(x)$, se se ne usano troppi, pochi dati andranno nei singoli bin e quindi si avranno grosse fluttuazioni percentuali e quindi errori.

Notare che solo raramente i dati istogrammati saranno indipendenti: in questo caso il numero di dati che entrano in un dato bin è distribuito approssimativamente secondo una distribuzione di Poisson con parametro $\mu = N \cdot p$, essendo p la probabilità che un valore vada nel bin i oggetto ed N il numero totale dei dati. Ricordiamo che la deviazione standard è $\sigma = \sqrt{\mu}$.

In genere i dati sono correlati e la situazione è peggiore.

Un altro modo per avere informazione sulla $f(x)$ è stimarne il valor medio (per esempio con la media aritmetica), la varianza ed eventualmente i momenti superiori. Anche in questo caso la convergenza statistica verso il valore vero è rallentata dalla presenza di correlazioni. Per esempio, per il caso del valor medio, supponiamo che il segnale abbia una autocovarianza $C(k)$ e vogliamo stimarne il valor medio facendo la media di n campioni successivi

$$(8.1) \quad \mu = \frac{1}{N} \sum_{i=1}^N x_i$$

e la varianza su questa stima è

$$(8.2) \quad \sigma_{\mu}^2 = \frac{1}{N} \sum_{k=-N}^N C(k) \cdot \left(1 - \frac{|k|}{N}\right)$$

Nel caso in cui

$$(8.3) \quad C(k) = \sigma^2 \cdot e^{-\frac{|k|}{\tau}}$$

abbiamo approssimativamente (per $\tau > 1$)

$$(8.4) \quad \sigma_{\mu}^2 \approx \frac{2 \cdot \tau \cdot \sigma^2}{N} \cdot \left(1 - \frac{1 - e^{-\frac{N}{\tau}}}{\frac{N}{\tau}} \right) = \frac{2 \cdot \sigma^2}{N'} \cdot \left(1 - \frac{1 - e^{-N'}}{N'} \right)$$

dove $N' = \frac{N}{\tau}$ è la “lunghezza equivalente” del pezzo di dati. Se N' è grande, si ha circa

$$(8.5) \quad \sigma_{\mu}^2 \approx \frac{2 \cdot \tau \cdot \sigma^2}{N} \cdot \left(1 - \frac{1 - e^{-\frac{N}{\tau}}}{\frac{N}{\tau}} \right) = \frac{2 \cdot \sigma^2}{N'} \cdot \left(1 - \frac{1 - e^{-N'}}{N'} \right)$$

Autocorrelazione

è che, come la definizione di autocorrelazione di un processo stocastico è data dalla (7.12), che nell'ipotesi stazionaria diventa

$$(8.6) \quad R(\tau) = E[X(t+\tau) \cdot X^*(t)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2^* \cdot f(x_1, x_2; \tau) \cdot dx_1 \cdot dx_2$$

Data la commutatività della moltiplicazione, la parte reale è una funzione pari, la parte immaginaria (se presente) è dispari.

Inoltre gode delle seguenti proprietà:

- a) $R(0)$ è reale ed è il massimo assoluto; possono esserci altri valori di τ in cui $R(\tau) = R(0)$, ma solo se la X è periodica e τ è un multiplo del periodo
- b) $R(\tau) = R^*(-\tau)$, quindi la sua trasformata di Fourier è reale
- c) La sua trasformata di Fourier è non-negativa. Questa proprietà deriva dalla proprietà di essere **definita positiva**, cionque presi n numeri arbitrari a_1, a_2, \dots, a_n ed n numeri reali $\tau_1, \tau_2, \dots, \tau_n$, si deve avere

$$(8.7) \quad \sum_{i=1}^n \sum_{l=1}^n R(\tau_i - \tau_l) \cdot a_i \cdot a_l \geq 0$$

Nel caso discreto la situazione è analoga, ma sostituiamo la variabile continua τ con la discreta k e la densità di probabilità con la distribuzione:

$$(8.8) \quad R(k) = E[X(i+k) \cdot X^*(i)] = \sum_{i_1=-\infty}^{\infty} \sum_{i_2=-\infty}^{\infty} x_{i_1} x_{i_2}^* \cdot f(x_{i_1}, x_{i_2}; k)$$

Data l'ipotesi di ergodicità, possiamo stimare la $R(k)$ sostituendo il valore atteso della (8.6) con la media su un pezzo di dati $\{x_1, x_2, \dots, x_N\}$:

$$(8.9) \quad R(k) = \frac{1}{N-k} \sum_{i=1}^{N-k} x_{i+k} \cdot x_i^*$$

Una diversa stima è

$$(8.10) \quad R(k) = \frac{1}{N} \sum_{i=1}^{N-k} x_{i+k} \cdot x_i^*$$

Come si vede, la differenza è nel coefficiente di normalizzazione. Il primo stimatore è non-distorto (unbiased), ma sui valori di k vicini a N ci sono forti incertezze (fluttuazioni).

La seconda stima è chiaramente distorta (il valore aspettato di $R(k)$ è inferiore al valor vero di un fattore $\frac{N}{N-k}$, ma così sono ridotte anche le fluttuazioni), ma la (8.10) gode della proprietà b).

Spesso nell'analisi dei segnali si usa la tecnica di moltiplicare i dati per una "finestra", cioè dei pesi opportuni, in genere minori alle estremità. Questa operazione viene detta "finestratura" (windowing).

La (8.10) si ricava dalla (8.9) tramite una "finestra" a triangolo isoscele detta "finestra di Bartlett".

Le procedure (8.9) e (8.10) sono usate raramente, a parte il caso in cui siamo interessati solo a un piccolo sottoinsieme dei ritardi k . Ciò perché il calcolo è molto pesante (abbiamo $N^2/2$ moltiplicazioni ed addizioni). Esse (o meglio la (8.10)) vengono sostituite dalla seguente procedura che si avvale della FFT (Fast Fourier Transform).

Consideriamo la successione infinita $\{x_i\}$ composta da tutti 0, a parte i valori per i da 1 a N , pari ai nostri campioni. La stima (8.10) è ottenuta tramite la

$$(8.11) \quad R(k) = \frac{1}{N} \sum_{i=-\infty}^{\infty} x_{i+k} \cdot x_i^*$$

Ricordiamo che la convoluzione tra le successioni a e b (eq. (6.8)) è

$$(8.12) \quad y_k = \sum_{i=-\infty}^{\infty} a_{k-i} \cdot b_i$$

e si vede quindi che la (8.11) è la convoluzione tra la x invertita temporalmente e la coniugata della x . Ne segue che, essendo $X(\omega)$ la trasformata di Fourier della successione x ,

$$(8.13) \quad R(k) = \frac{1}{N} \cdot \mathcal{F}^{-1}(|X(\omega)|^2)$$

dove con \mathcal{F}^{-1} indichiamo la trasformata inversa di Fourier. Per eseguire questa procedura ci si avvale della trasformata di Fourier (per i dettagli, si veda al capitolo sui filtri in frequenza).

Spettro di potenza

Lo spettro di potenza è definito come la trasformata di Fourier dell'autocorrelazione (vedi eq.(7.17)). Nel caso stazionario si ha

$$(8.14) \quad S(\omega) = \int_{-\infty}^{\infty} R(\tau) \cdot e^{-j\omega\tau} d\tau$$

e, date le proprietà dell'autocorrelazione, è reale e non-negativo. Se inoltre il processo è reale, lo spettro è una funzione pari.

Si noti che le proprietà dello spettro e dell'autocorrelazione sono analoghe a quelle della densità di probabilità e della funzione caratteristica.

Mentre l'autocorrelazione deve avere le proprietà ("complicate") indicate nel paragrafo precedente, lo spettro ha l'unica limitazione di essere positivo (o meglio non negativo). In genere comunque data una funzione non negativa di ω , a singolo valore, possiamo costruire un processo stocastico che la abbia come spettro di potenza.

Per i processi discreti, abbiamo (eq. (7.69))

$$(8.15) \quad S(\Omega) = \sum_{i=-\infty}^{\infty} R(i) \cdot e^{-j\Omega i}$$

essendo Ω la pulsazione normalizzata; ricordiamo che per avere lo spettro nelle unità "fisiche", $\Omega = \omega \cdot \Delta t$, e

$$(8.16) \quad S'(\omega) = S(\Omega) \cdot \Delta t$$

Se il segnale è reale, la potenza del segnale nella banda di frequenza $\{\nu_1, \nu_2\}$ è data da

$$(8.17) \quad 2 \cdot \int_{\nu_1}^{\nu_2} S'(2\pi \cdot \nu) d\nu$$

Il fattore 2 deriva dal fatto che per segnali reali non c'è differenza tra frequenze positive e negative.

Se il segnale è complesso, $S'(\omega)$ non è pari e bisogna distinguere tra frequenze positive e negative e quindi $\int_{\nu_1}^{\nu_2} S'(2\pi \cdot \nu) d\nu$ è diverso da $\int_{-\nu_2}^{-\nu_1} S'(2\pi \cdot \nu) d\nu$. Perciò la potenza del segnale nella banda di frequenza $\{\nu_1, \nu_2\}$ è data da

$$(8.18) \quad \int_{\nu_1}^{\nu_2} S'(2\pi \cdot \nu) d\nu$$

Vogliamo notare che spesso nella fisica sperimentale e nella tecnica, quando i segnali sono sempre reali e non interessano molto le questioni teoriche, si usa una diversa definizione di spettro, con le sole frequenze positive e con un valore doppio di quello ottenuto da (7.69). Questo viene chiamato **spettro unilatero**. Scompare così lo “scomodo” fattore 2 della (8.17).

Vediamo ora come stimare lo spettro di potenza.

Stimatori spettrali non parametrici

Un modo per stimare lo spettro è quello di stimare l'autocorrelazione con lo stimatore (8.10), cioè lo stimatore distorto con la finestra di Bartlett applicata e farne la trasformata di Fourier.

(Se si usasse la stima (8.9), avremmo una stima spettrale più “scomoda”, con errori più grandi e, per certe frequenze, con valori negativi. Ciò è dovuto al grosso peso che hanno le fluttuazioni sui valori estremali).

Questa procedura tuttavia non è usata quasi mai. essa infatti è computazionalmente molto pesante. Lo stesso risultato lo si ottiene usando la procedura usata per il calcolo veloce dell'autocorrelazione (escludendo la trasformata inversa). La ripetiamo:

Siano dati N campioni successivi $\{x_1, x_2, \dots, x_N\}$. Se ne faccia la trasformata di Fourier (per dati discreti)

$$(8.19) \quad X(\Omega) = \sum_{i=1}^N x_i \cdot e^{-j(i-1)\Omega}$$

La stima spettrale è data da

$$(8.20) \quad S(\Omega) = \frac{|X(\Omega)|^2}{N}$$

È questa una conseguenza del teorema di Wiener-Kinchin²². La trasformata (8.19) viene eseguita in modo veloce con la FFT.

Se si vogliono usare le frequenze “fisiche”, con tempo di campionamento Δt , ricordiamo che $\Omega = \omega \cdot \Delta t$ e

$$(8.21) \quad S(\omega) = \frac{|X(\omega \cdot \Delta t)|^2}{N} \cdot \Delta t$$

Lo spettro (e la sua stima) è una grandezza statistica: da esso ovviamente non possono ricavarsi i dati originari e ciò perché si è persa l'informazione di fase delle componenti armoniche.

Vediamo tre problemi dello stimatore (8.20):

- **La risoluzione spettrale**

Per come è costruita la stima spettrale, notiamo che la risoluzione in frequenza della stima spettrale dipende dalla lunghezza del pezzo di dati. Se abbiamo N campioni, essa è, convenzionalmente, in unità della frequenza fisica,

$$(8.22) \quad \Delta \nu = \frac{1}{N \cdot \Delta t}$$

Discuteremo meglio questo argomento quando avremo introdotto il concetto di finestra.

Gli stimatori parametrici non hanno questo problema (ma spesso non possono essere usati).

- **Problema del finestramento**

²² In effetti in passato la (8.20) definiva lo spettro di potenza e il teorema di Wiener-Kinchin (noto già, precedentemente, ad Einstein in una forma un po' diversa) stabiliva che la stessa informazione si aveva facendo la trasformata di Fourier dell'autocorrelazione.

La stima è fatta su un pezzo di dati non infinito (come dovrebbe essere in teoria), ma di lunghezza finita. Questo può essere visto come il prodotto della successione infinita per una finestra “rettangolare” (boxcar in Inglese)²³

$$(8.23) \quad w_i = \begin{cases} 1 & \text{per } 1 \leq i \leq N \\ 0 & \text{altrove} \end{cases}$$

Ora sappiamo che la trasformata di Fourier del prodotto di due funzioni è pari alla convoluzione delle trasformate delle due funzioni, quindi la stima spettrale data dalla (8.20) “smussa” lo spettro vero: cioè è come se ci si passasse sopra un passa basso simile a quello di equazione (6.56) (in effetti la risposta in fase è diversa).

Questo smussamento può essere particolarmente fastidioso (se si vuole stimare spettri con rapide variazioni). Si riduce il problema usando una finestra non rettangolare.

Lo stimatore è quindi così costruito:

$$(8.24) \quad S(\Omega) = \frac{\left| \sum_{i=1}^N w_i \cdot x_i \cdot e^{-j(i-1)\Omega} \right|^2}{\sum_{i=1}^N |w_i|^2}$$

La successione $\{w_i\}$ è la “finestra”. In genere vale 1 (il valore massimo) per i valori centrali, è simmetrica e si riduce agli estremi.

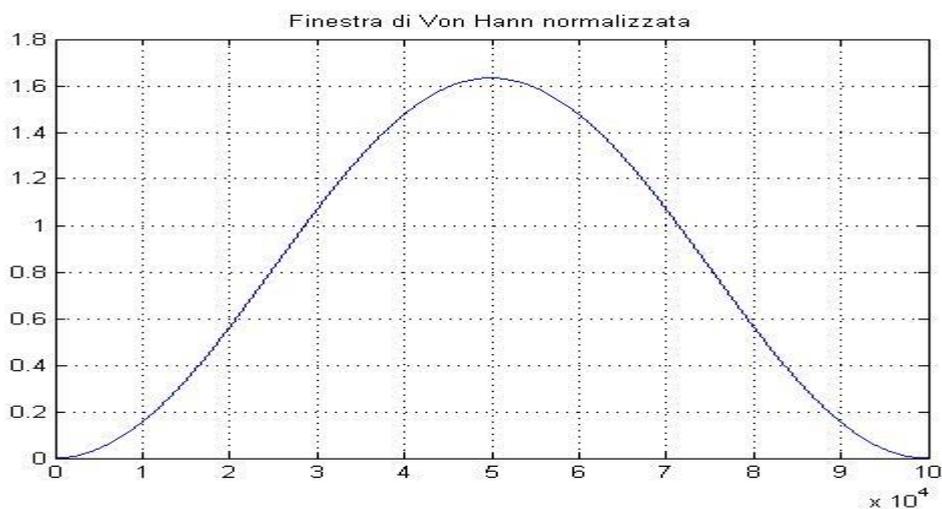
Una finestra molto usata è la finestra che prende il nome da Von Hann²⁴ (detta anche *hanning window*) è la seguente

$$(8.25) \quad w_i = \frac{1 + \cos\left(\frac{2\pi \cdot (i - N/2)}{N}\right)}{2}$$

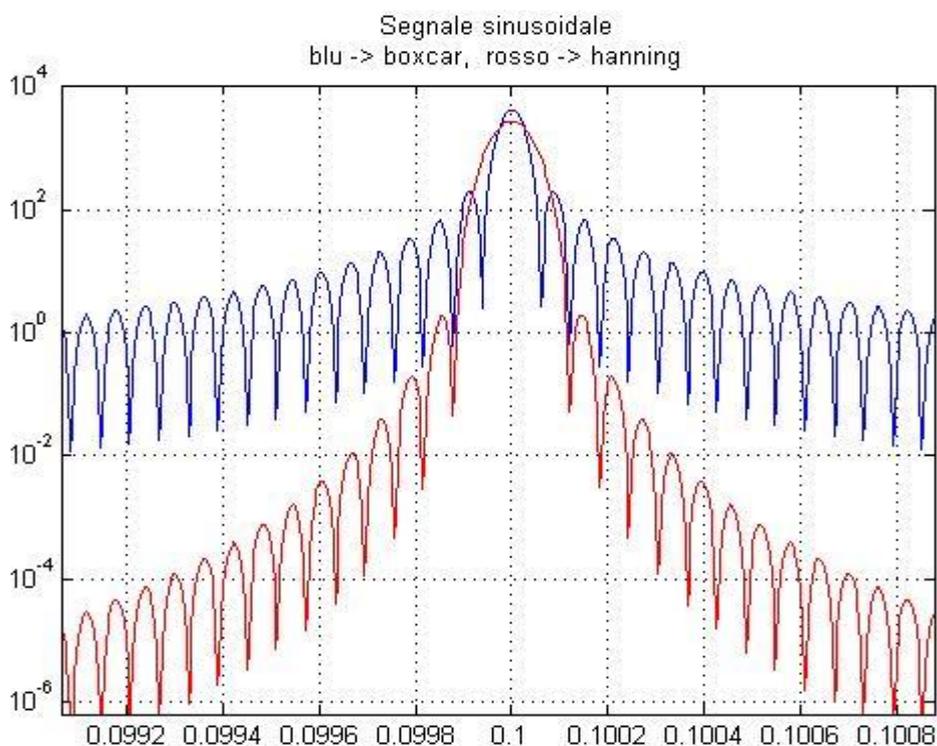
²³ La presenza di questa finestra sui dati temporali implica la finestra di Bartlett sull'autocorrelazione: l'autocorrelazione di una funzione rettangolare è una funzione triangolare isoscele.

²⁴ Julius Von Hann (1839-1921), metereologo austriaco. Blackmann e Tukey hanno dato il suo nome a questa finestra.

Ecco la finestra di Von Hann normalizzata per $\sum_{i=1}^N |w_i|^2$ (100000 punti):



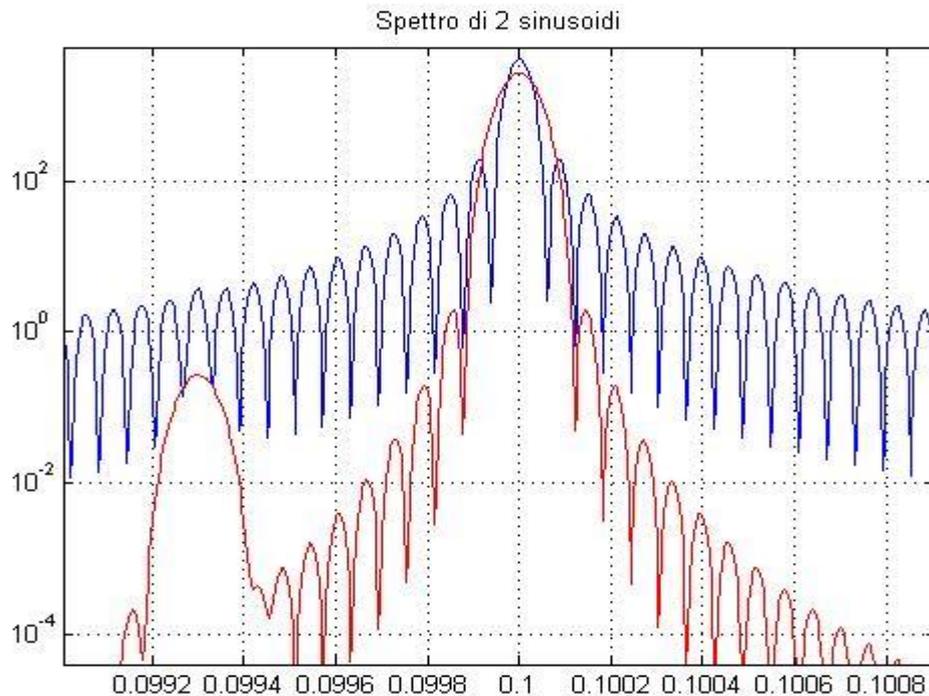
Vediamo ora l'effetto della finestra. Nel grafico sono riportate le stime spettrali per il caso di una sinusoidale di frequenza 0.1, eseguite con la finestra di hanning e con la finestra rettangolare (o "boxcar", cioè' usando semplicemente la (8.10)):



Lo spettro "ideale" del segnale sarebbe una delta a frequenza 0.1. In effetti invece vediamo un picco con dei lobi laterali (artefatti causati dalla finestratura dei dati).

Notiamo che l'aver introdotto la finestra di Von Hann ha ridotto la risoluzione (di circa un fattore 2; ma questo in genere non è un problema, infatti se occorre una migliore risoluzione basta prendere un pezzo più lungo di dati), ma ha ridotto di parecchio l'ampiezza dei lobi laterali.

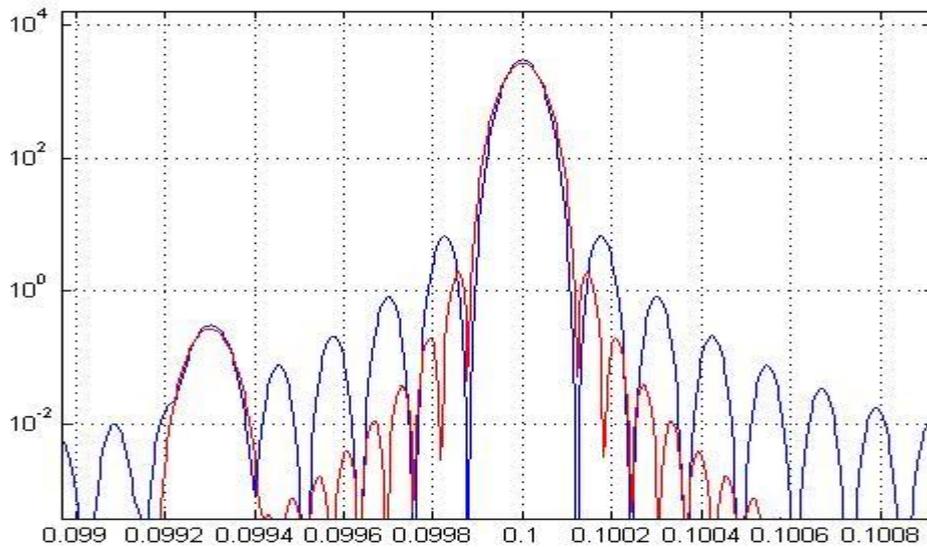
Nella figura seguente rappresentiamo le stime spettrali con le due finestre di un segnale composto di due sinusoidi (una 100 volte più piccola dell'altra)



Si nota che i lobi laterali della finestra rettangolare nascondono completamente il picco più piccolo, perfettamente visibile invece con la finestra di Von Hann.

Sono state studiate numerose finestre, con proprietà diverse. Quale utilizzare dipende dal tipo di dati che si hanno.

Un ultimo esempio: la finestra triangolare (confrontata con quella di hanning)



Si noti che i lobi laterali sono più ampi, ma c'è una risoluzione (ed un'ampiezza di picco) leggermente migliore. I lobi poi hanno una larghezza doppia.

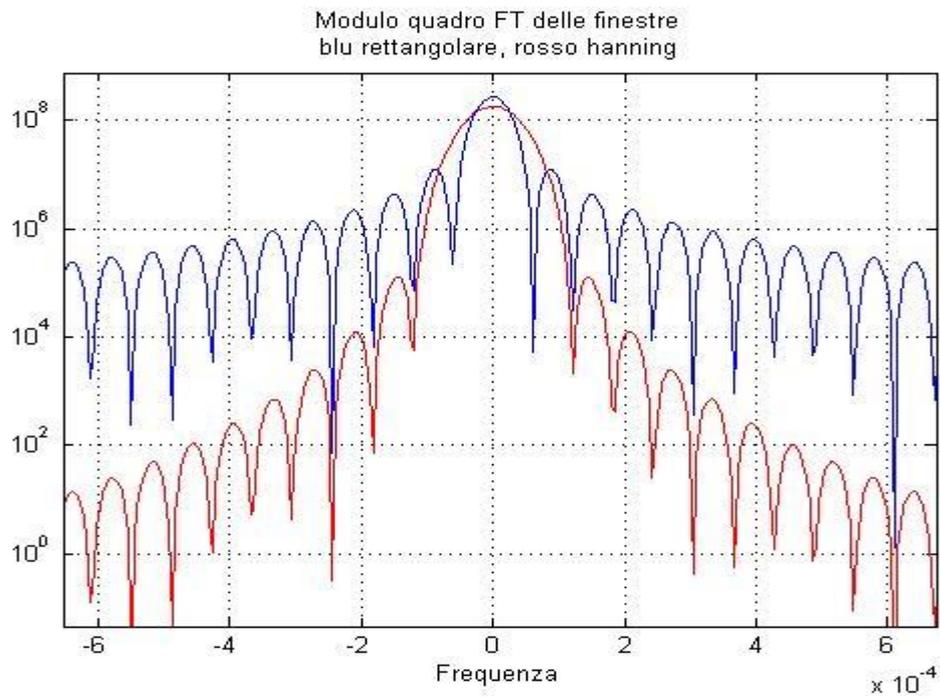
Infine torniamo al concetto di risoluzione: introdurre una finestra (anche la rettangolare, che spesso è una "finestra involontaria") significa "smussare" lo spettro col modulo quadro della trasformata di Fourier della finestra. Se lo spettro è $S(\omega)$ e se la finestra $w(t)$ ha trasformata di Fourier $W(\omega)$, il valore atteso dello spettro "finestrato" è

$$(8.26) \quad S_w(\omega) = k \cdot S(\omega) \otimes |W(\omega)|^2$$

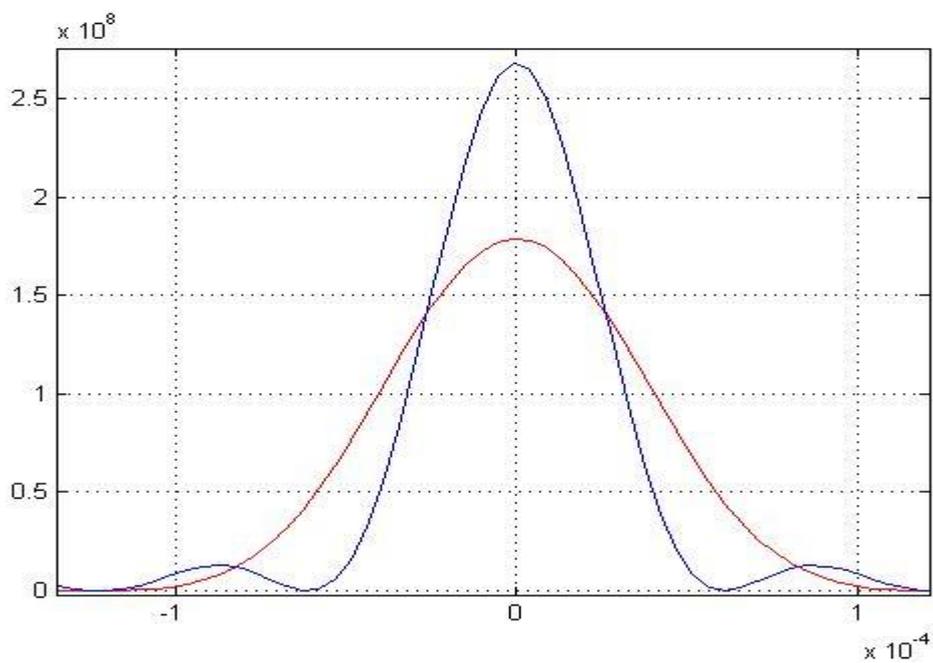
dove k è una costante di normalizzazione.

Se prendiamo il segnale di più alta risoluzione possibile nello spettro, una delta spettrale, cioè una sinusoide nel tempo, possiamo definire la risoluzione come la larghezza con cui è visto questo segnale (per esempio la larghezza a mezza altezza; definizioni migliori per la risoluzione possono farsi, per esempio, conoscendo il rapporto segnale-rumore). Questo non è approssimativamente che il modulo quadro della trasformata di Fourier della finestra (a parte una costante moltiplicativa e una traslazione).

Ecco il grafico, in scala logaritmica, per le finestre rettangolare e hanning



ed ecco lo stesso grafico in scala lineare



Da questo vediamo che la risoluzione (col metodo della larghezza a mezza altezza) è circa 1.5 volte peggiore nell'hanning.

- **Incertezza della stima**

Un problema dello stimatore (8.20) o di quello (8.24) è che l'incertezza della stima è molto elevata, e, soprattutto, non si riduce con l'aumentare della lunghezza del pezzo di dati.

Per evidenziare la cosa, supponiamo che il nostro segnale sia una successione scorrelata di campioni gaussiani normalizzati, cioè un rumore bianco con densità spettrale $S(\Omega) = 1$.

Per semplicità usiamo lo stimatore (8.20). La trasformata di Fourier è data da

$$(8.27) \quad X(\Omega) = \sum_{i=1}^N x_i \cdot e^{-j(i-1)\Omega} = \sum_{i=1}^N x_i \cdot \cos((i-1)\cdot\Omega) + j \cdot \sum_{i=1}^N x_i \cdot \sin((i-1)\cdot\Omega)$$

Consideriamo il pezzo $\sum_{i=1}^N x_i \cdot \cos((i-1)\cdot\Omega)$. Per ogni valore di Ω abbiamo la somma di

N numeri casuali $x_i \cdot \cos((i-1)\cdot\Omega)$ che hanno media 0 e varianza $\frac{1}{2}$. Per il teorema del

limite centrale quindi il valore della sommatoria è una variabile normale con media

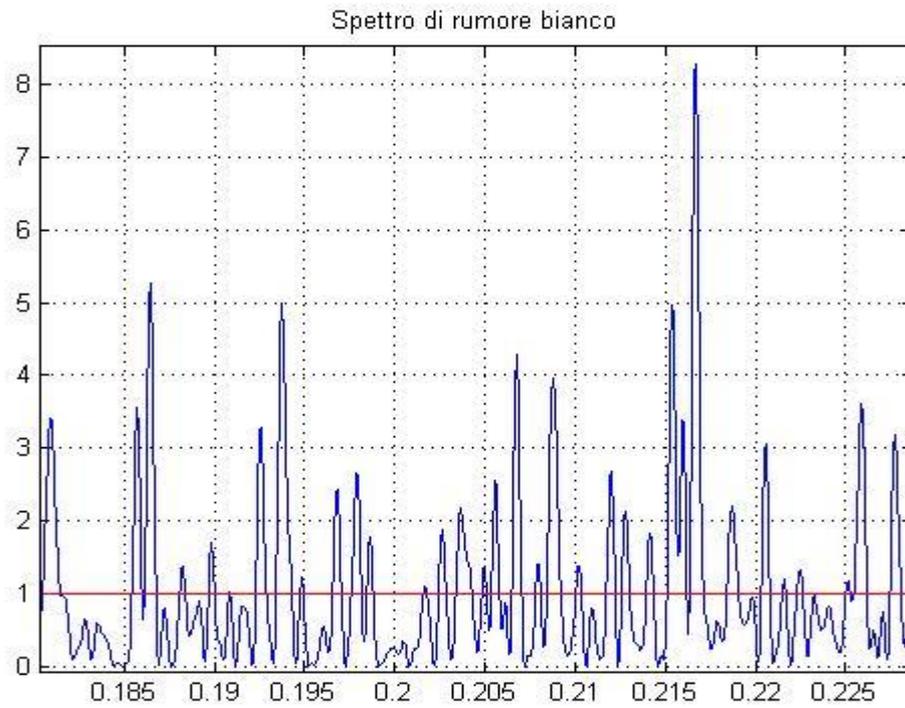
nulla e varianza $\frac{N}{2}$. Lo stesso può dirsi dell'altra sommatoria con i seni. $|X(\Omega)|^2$ è una

variabile casuale ottenuta dalla somma dei quadrati di due variabili gaussiane: ha quindi distribuzione esponenziale²⁵, con valore atteso N . Poi, poiché per la stima dello spettro, il modulo quadro della trasformata di Fourier va divisa per N , abbiamo che **la distribuzione dei valori dello spettro è esponenziale con media 1**. (Ricordiamo che ci aspetteremmo semplicemente 1).

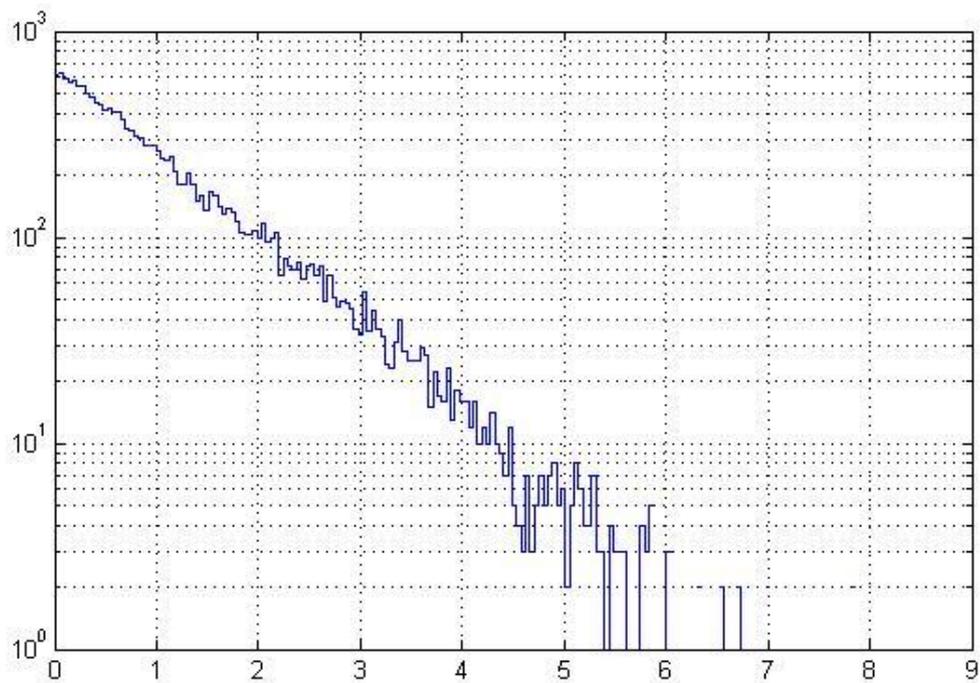
Ecco un pezzo dello spettro di questo genere

²⁵ La distribuzione esponenziale, a parte un coefficiente di normalizzazione, è la distribuzione del χ^2 con 2 gradi di libertà. In essa la deviazione standard è uguale al valore atteso. La

probabilità di avere valori superiori di x è $P = e^{-\frac{x}{\mu}}$.



Si noti che ci aspetteremmo una linea costante a 1 (linea rossa). L'istogramma delle ampiezze spettrali è



Si noti che la distribuzione è esponenziale (una retta in plot semilog).

Attenzione ! la situazione non cambia minimamente con l'aumentare del numero di campioni.

Col numero dei campioni invece aumenta il numero di punti indipendenti dello spettro (la risoluzione).

Notiamo inoltre che la distribuzione esponenziale ha code abbastanza pesanti, cioè valori di probabilità elevate anche per alti rapporti critici²⁶. Per esempio, per CR=3 abbiamo $P(x > 4) = e^{-4} = 0.0183$.

Per ridurre quindi le fluttuazioni, possiamo operare in due modi:

- smussare la stima spettrale con un opportuno filtro passa-basso
- ricavare varie stime spettrali del tipo (8.20) o (8.24) da pezzi di dati indipendenti e farne la media.

Quest'ultima è la procedura più usata (in genere è più comodo fare FFT più corte).

Avendo a disposizione un pezzo di $M \cdot N$ campioni, lo stimatore più generale è quindi

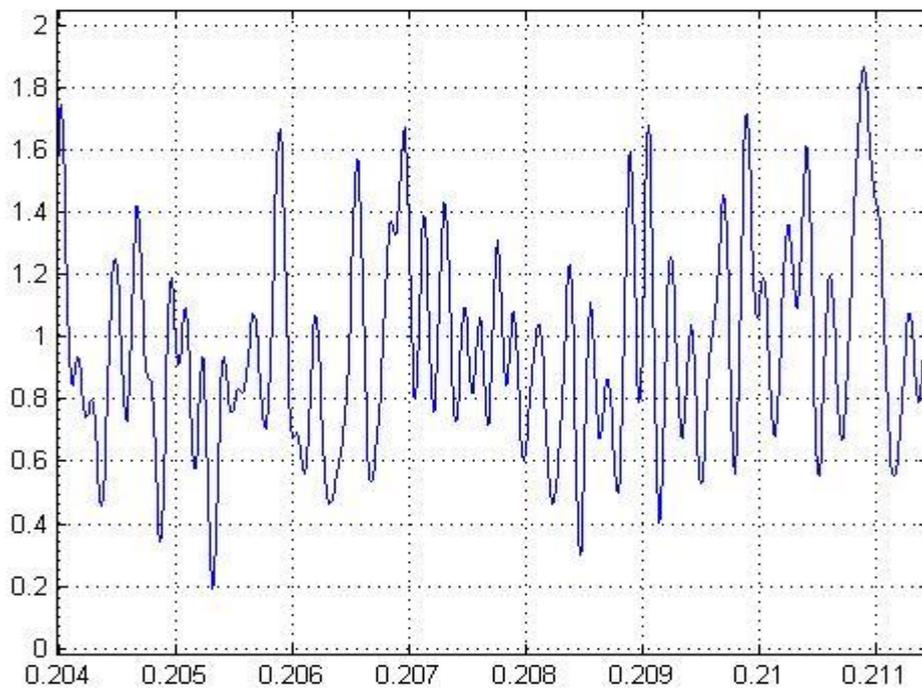
$$(8.28) \quad S(\Omega) = \frac{1}{M} \cdot \sum_{k=1}^M \frac{\left| \sum_{i=1}^N w_i \cdot x_{i+(k-1) \cdot N} \cdot e^{-j(i-1)\Omega} \right|^2}{\sum_{i=1}^N |w_i|^2}$$

La scelta di M (il numero di pezzi) si fa mediando tra la necessità di una buona risoluzione e quella di una bassa incertezza.

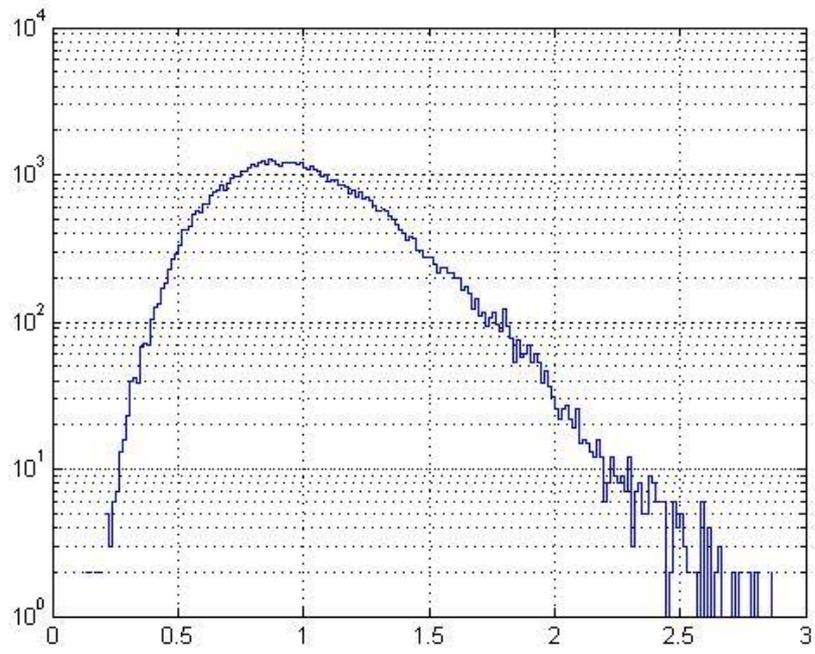
²⁶ Ricordiamo che il rapporto critico (critical ratio) è definito come $CR = \frac{|x - \mu|}{\sigma}$.

Nell'ipotesi che abbiamo fatto in precedenza, di rumore bianco e finestra rettangolare, la distribuzione attesa è una distribuzione del χ^2 normalizzata (a valore atteso 1) con $2M$ gradi di libertà.

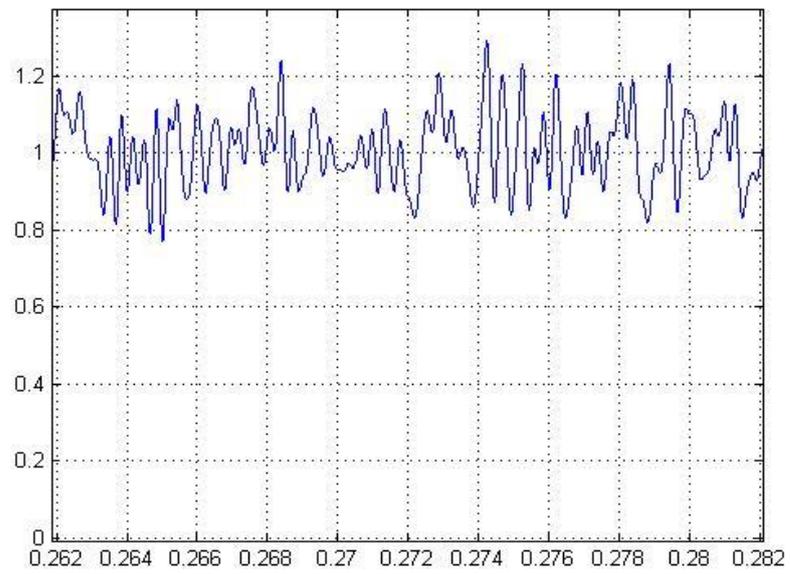
Ecco il grafico di un pezzo di spettro stimato con finestra rettangolare ed $M=10$.



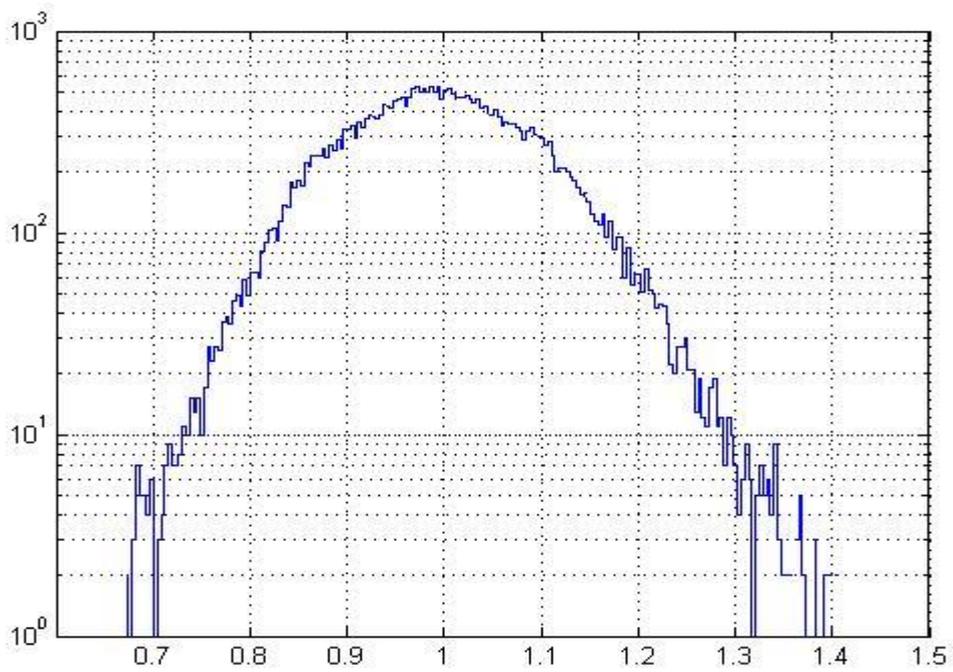
ed ecco l'istogramma



Ed ecco con $M=100$:



e la distribuzione è



Nel caso di finestre non rettangolari, è ragionevole che i pezzi siano interallacciati. Infatti le parti iniziali e finali di ciascun pezzo contribuiscono molto poco al valore spettrale: l'interallacciamento (per esempio a metà) fa sì che tutti i dati siano considerati.

Stimatori spettrali parametrici

Un modo alternativo di stimare lo spettro di potenza di un segnale è scegliere un modello che riteniamo adeguato e stimarne i parametri, quindi calcolare lo spettro di potenza del modello con quei parametri.

Uno dei modelli più usati è il modello AR, detto anche “a massima entropia”, perché presuppone un andamento asintotico esponenziale dell'autocorrelazione, che corrisponde alla massima entropia.

Facciamo un esempio:

- riteniamo che un certo segnale che stiamo misurando sia adeguatamente modellizzato da un processo AR di un certo ordine:

$$(8.29) \quad y_i = b_0 \cdot x_i + w_1 \cdot y_{i-1} + w_2 \cdot y_{i-2} \dots + w_n \cdot y_{i-n}$$

Ora moltiplichiamo il membro a destra e a sinistra per y_{i-k} , con $k > 0$, e prendiamone il valore atteso. Poiché $R_{xy}(k) = 0$ per $k > 0$, si ha

$$(8.30) \quad R_{yy}(k) = w_1 \cdot R_{yy}(k-1) + w_2 \cdot R_{yy}(k-2) \dots + w_n \cdot R_{yy}(k-n)$$

e, prendendo tutti i $1 \leq k \leq n$ e indicando $r_k = R_{yy}(k)$, abbiamo

$$(8.31) \quad \begin{pmatrix} r_0 & r_1 & \dots & r_{n-1} \\ r_1 & r_0 & \dots & r_{n-2} \\ \vdots & \vdots & & \vdots \\ r_{n-1} & r_{n-2} & \dots & r_0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$$

Queste sono le equazioni di Yule e Walker. Con esse si possono ricavare i coefficienti del modello AR dai valori dell'autocorrelazione.

La matrice simmetrica che compare viene detta matrice di Toeplitz e il sistema viene risolto con l'efficiente algoritmo di Levinson e Durbin.

Se nelle (8.31) mettiamo, al posto dei valori dell'autocorrelazione, delle loro stime, possiamo stimare i valori dei parametri w_i del modello AR(n). Da questi lo spettro viene stimato (vedi (6.30)) come

$$(8.32) \quad S_y(\Omega) = \sigma_x^2 \cdot \left| F(e^{j\Omega}) \right|^2 = \frac{b_0^2 \cdot \sigma_x^2}{\left| 1 - \sum_{k=1}^n w_k \cdot e^{-j\Omega k} \right|^2}$$

Questo tipo di stima spettrale viene implementata in vari modi, a seconda di come viene eseguita la stima dell'autocorrelazione o risolto il sistema (8.31).

Gli stimatori spettrali parametrici sono particolarmente buoni ed efficienti se il modello usato è a) aderente ai dati, b) non ha un numero troppo elevato di parametri.

Cross-correlazione e cross-spettro

La correlazione incrociata (o cross-correlazione o anche intercorrelazione) tra due processi stocastici stazionari X e Y è data da

$$(8.33) \quad R_{XY}(\tau) = E[X(t+\tau) \cdot Y^*(t)]$$

La trasformata di Fourier di $R_{XY}(\tau)$, indicata in genere con $S_{xy}(\omega)$, è detta **cross-spettro** o **spettro incrociato di potenza**.

Una funzione di autocorrelazione non ha le simmetrie e limitazioni di forma dell'autocorrelazione. Così il cross-spettro può assumere valori negativi o anche complessi.

Nell'ipotesi di ergodicità, possiamo stimare la $R_{XY}(\tau)$ sostituendo il valore aspettato della (8.33) con la media su due pezzi di dati $\{x_1, x_2, \dots, x_N\}$ e $\{y_1, y_2, \dots, y_N\}$

$$(8.34) \quad R_{xy}(k) = \frac{1}{N-k} \sum_{i=1}^{N-k} x_{i+k} \cdot y_i^*$$

Una diversa stima è

$$(8.35) \quad R_{xy}(k) = \frac{1}{N} \sum_{i=1}^{N-k} x_{i+k} \cdot y_i^*$$

Per quanto riguarda la stima del cross-spettro, questa può ottenersi come trasformata di Fourier di una stima della cross-correlazione, o tramite l'uso del teorema di Wiener-Khinchin, per esempio nel modo seguente.

Si dividono i dati disponibili per ciascun processo in M pezzi di N dati. Si calcola quindi

$$(8.36) \quad S_{xy}(\Omega) = \frac{1}{M} \cdot \sum_{k=1}^M \frac{\left(\sum_{i=1}^N w_i \cdot x_{i+(k-1) \cdot N} \cdot e^{-j(i-1)\Omega} \right) \cdot \left(\sum_{i=1}^N w_i \cdot y_{i+(k-1) \cdot N}^* \cdot e^{-j(i-1)\Omega} \right)}{\sum_{i=1}^N |w_i|^2}$$

La scelta di M (il numero di pezzi) si fa mediando tra la necessità di una buona risoluzione e quella di una bassa incertezza.

Coerenza

Lo spettro di potenza incrociato, come la cross-correlazione, è una misura della similarità di due processi. La **coerenza** (coherence) è una forma normalizzata del cross-spettro. La possiamo definire²⁷ come

$$(8.37) \quad C_{xy}(\Omega) = \frac{S_{xy}(\Omega)}{\sqrt{S_x(\Omega) \cdot S_y(\Omega)}}$$

La stima della coerenza si fa con

$$(8.38) \quad C_{xy}(\Omega) = \frac{S_{xy}(\Omega)}{\sqrt{S_x(\Omega) \cdot S_y(\Omega)}}$$

Attenzione però. Le stime vanno fatte con equazioni come la (8.36) e la (8.28), con un valore di M abbastanza elevato. Per M=1 la (8.38) dà 1.

²⁷ Talvolta la coerenza viene definita come il quadrato della presente definizione.

9 - Filtraggio

Live Script

Segnali e rumori, rapporto segnale/rumore

Il tipico problema di misura è il seguente:

Siamo interessati a stimare un segnale $x(t)$ (o la sua trasformata $X(\omega)$), ma possiamo osservare $y(t)$ che

- contiene una trasformazione x' di x
- contiene addizionato un altro segnale non desiderato

Supporremo che la trasformazione del segnale x sia lineare e descritta da un sistema F , e che il segnale non desiderato aggiuntivo (che chiamiamo **rumore**) sia un processo gaussiano descritto da uno spettro

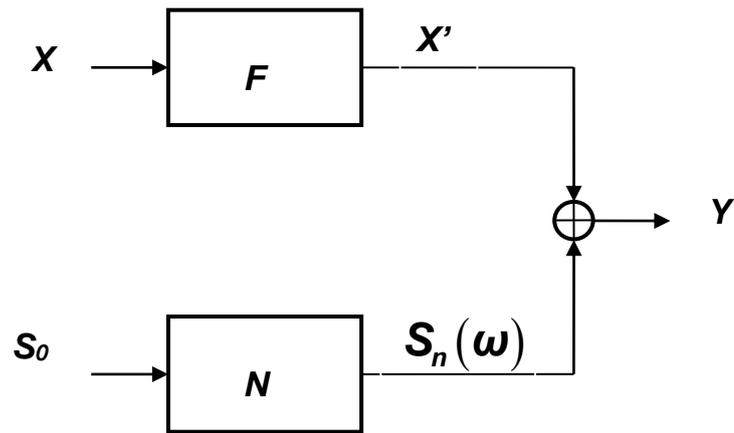
$$(9.1) \quad S_n(\omega) = S_0 \cdot |N(\omega)|^2$$

dove S_0 è una costante e N è uno degli infiniti sistemi che risolvono l'equazione (9.1)

La situazione è descritta, nel dominio trasformato, dall'equazione

$$(9.2) \quad Y(\omega) = X(\omega) \cdot F(\omega) + S_0 \cdot N(\omega)$$

Mostriamo lo schema di questo processo



Nell'ipotesi che il rumore non esista (o sia trascurabile) il problema si riduce a trovare il sistema inverso di F (se questo è invertibile) o una sua buona approssimazione.

Più complessa (e praticamente interessante) è il caso in cui il rumore non sia trascurabile.

Distinguiamo tre casi particolarmente interessanti:

1. il segnale $x(t)$ ha una forma nota,

$$(9.3) \quad x(t - t_0) = A \cdot w(t)$$

e va individuata la sua presenza, e stimati il tempo di occorrenza t_0 e l'ampiezza A .

Questo è il problema della **rivelazione (detection)**. In tal caso si definisce **rapporto segnale/rumore (SNR - signal-to-noise ratio)** il rapporto tra il massimo del segnale x e la deviazione standard del rumore

$$(9.4) \quad SNR = \frac{\max(x(t))}{\sigma_n}$$

2. il segnale $x(t)$ è un processo stocastico che vogliamo stimare. In questo caso possiamo definire SNR il rapporto tra la deviazione standard del segnale e quella del rumore²⁸

(9.5)
$$SNR = \frac{\sigma_x}{\sigma_n}$$

3. il segnale $x(t)$ è un processo stocastico di cui vogliamo stimare uno o più parametri (per esempio il valor medio e/o la varianza). In questo caso ci interessa l'incertezza sulla stima del parametro. L'SNR può essere in questo caso l'incertezza relativa (quando questa ha senso).

In tutti questi casi ci interessa di ridurre l'effetto del rumore. Un altro modo di dire ciò è "aumentare il rapporto segnale/rumore"²⁹. Ciò però ha senso se la distribuzione del rumore rimane la stessa dopo la stima (per esempio è e rimane gaussiana). Altrimenti il SNR non è un buon indice, ciò che conta sono le probabilità, per esempio di avere un falso segnale o di avere un certo errore di stima o di farsi "sfuggire" segnali non abbastanza ampi.

Problemi come questi si risolvono con la costruzione di particolari sistemi detti **filtri**, con particolari caratteristiche. Vedremo in questo capitolo come calcolare i migliori filtri per un dato scopo e come realizzarli praticamente.

²⁸ Talora si usa come SNR il rapporto delle varianze, cioè il quadrato di questa quantità. Definiremo tale grandezza **rapporto segnale/rumore di potenza (power SNR)**

²⁹ Ovviamente non basta ridurre il rumore: ciò si avrebbe, per esempio, semplicemente moltiplicando i dati per una costante piccola, ma ridurre il rumore non sacrificando l'ampiezza del segnale.

Il filtro adattato

Tratteremo qui il primo caso elencato e cioè il caso in cui il segnale $x(t)$ ha una forma nota $w(t)$,

$$(9.6) \quad x(t+t_0) = A \cdot w(t)$$

e va individuata la sua presenza, e stimati il tempo di occorrenza t_0 e l'ampiezza A . La soluzione di questo problema viene chiamata **filtro adattato (matched filter)**. Si noti che ottimizzare la stima di A significa anche ottimizzare il rapporto segnale/rumore, ma non viceversa (l'ottimizzazione dell'SNR lascia libero un parametro di ampiezza).

Supponiamo inoltre che $F=1$, cioè $x'=x$, per cui

$$(9.7) \quad y(t) = A \cdot w(t-t_0) + \varepsilon(t)$$

dove $\varepsilon(t)$ è il rumore.

Possiamo scrivere ciò in forma discretizzata

$$(9.8) \quad y_i = A \cdot w_{i-i_0} + \varepsilon_i$$

e discutere direttamente questo caso.

Ovviamente per definire la forma w abbiamo infinite scelte (ognuna con una diversa ampiezza); scegliamo quella che "normalizza" l'energia

$$(9.9) \quad \sum_{k=-\infty}^{\infty} w_k^2 = 1$$

Caso del rumore bianco

Supponiamo ora che il rumore sia bianco, quindi gli ε_i sono campioni gaussiani con spettro $S_n(\omega) = \sigma_\varepsilon^2$.

Cerchiamo il filtro migliore con la tecnica della massima verosimiglianza.

Possiamo scrivere la verosimiglianza ad un dato istante (per semplicità di notazione abbiamo scelto l'istante 0) come

$$(9.10) \quad L(A; \{y_i\}) = \prod_i \frac{1}{\sqrt{2\pi} \cdot \sigma_\varepsilon} \cdot e^{-\frac{(y_i - A \cdot w_i)^2}{2\sigma_\varepsilon^2}}$$

e prendendone il logaritmo ed ignorandone la parte costante abbiamo

$$(9.11) \quad l(A; \{y_i\}) = \sum_i -\frac{(y_i - A \cdot w_i)^2}{2\sigma_\varepsilon^2}$$

Eguagliando la derivata a 0 e sostituendo \hat{A} per A, si ha

$$(9.12) \quad \sum_i A \cdot w_i^2 = \sum_i y_i \cdot w_i$$

$$\hat{A} = \frac{\sum_i y_i \cdot w_i}{\sum_i w_i^2} = \sum_i y_i \cdot w_i$$

quindi, per ogni istante i,

$$(9.13) \quad \hat{A}_i = y_i \otimes w_{-k} = \sum_{k=-\infty}^{\infty} y_{i+k} \cdot w_k = \sum_{k=-\infty}^{\infty} (A \cdot w_{i+k-i_0} + \varepsilon_{i+k}) \cdot w_k$$

Cioè la stima di A è la cross-correlazione dei dati y per la forma del segnale w.

Poiché il problema è gaussiano (le statistiche coinvolte sono per ipotesi tutte normali), ci aspettavamo che la soluzione sarebbe stata un filtro lineare, come in effetti è.

Se non è presente alcun segnale, avremo in uscita un nuovo rumore

$$(9.14) \quad \eta_i = \sum_{k=-\infty}^{\infty} w_k \varepsilon_{i+k}$$

stazionario con varianza

$$(9.15) \quad \sigma_{\eta}^2 = \sigma_{\varepsilon}^2 \cdot \sum_{k=-\infty}^{\infty} w_k^2 = \sigma_{\varepsilon}^2$$

Se ci fosse solo il segnale, avremmo

$$(9.16) \quad \hat{A}_i = \sum_{k=-\infty}^{\infty} y_{i+k} \cdot w_k = A \cdot \sum_{k=-\infty}^{\infty} w_{i+k-i_0} \cdot w_k = A \cdot R_w(i-i_0)$$

$R_w(l)$ è l'autocorrelazione del segnale w_i , definita come

$$(9.17) \quad R_w(l) = \sum_{k=-\infty}^{\infty} w_{l+k} \cdot w_k = w_k \otimes w_{-k}$$

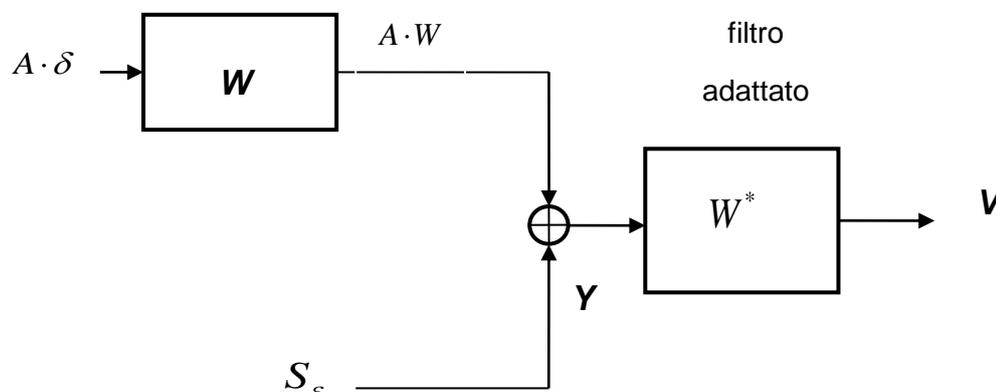
Se vogliamo un filtro che massimizzi il SNR, allora possiamo fare la correlazione tra i dati e $B \cdot w_k$, dove B è una qualsiasi costante, e avremo lo stesso risultato. Ma porre B=1 semplifica la valutazione dell'ampiezza del segnale.

Vediamo la soluzione nel dominio della frequenza.

Indichiamo l'uscita del filtro adattato con $\{v_i\}$, con trasformata V. Abbiamo

$$(9.18) \quad V(\omega) = Y(\omega) \cdot W^*(\omega)$$

Possiamo vederne lo schema:



Nel nodo somma vediamo la somma del segnale w (quando e se presente la delta all'ingresso di W) e del rumore bianco di densità spettrale S_ε .

La risposta del filtro adattato al segnale è pari alla risposta alla delta del sistema con funzione di trasferimento $|W(j\omega)|^2$,

Se i segnali w sono rari (o assenti), $\{v_i\}$ è la risposta del filtro adattato al rumore bianco; avrà quindi spettro

$$(9.19) \quad S_w(\omega) = S_\varepsilon \cdot |W(j\omega)|^2$$

e autocorrelazione pari all'autocorrelazione del segnale w moltiplicato per σ_ε^2 . Quindi lo spettro del rumore e la funzione di trasferimento (sempre reale), ovvero l'autocorrelazione del rumore e la risposta al segnale, hanno la stessa forma. Questa è un'interessante proprietà del filtro adattato.

C'è stato miglioramento di SNR? Dato che la varianza del rumore è la stessa prima e dopo, vediamo cosa succede sul segnale. Prima il massimo valore era

$$(9.20) \quad Max_{pre} = A \cdot \max(|w_i|)$$

dopo diventa

$$(9.21) \quad \text{Max}_{post} = A$$

Ora è evidente che, per la (9.9), $\max(|w_i|) \leq 1$ (l'eguaglianza si ha se w_i è una delta). E il guadagno del filtro è

$$(9.22) \quad G = \frac{1}{\max(|w_i|)}$$

Altre dimostrazioni

Dato il grande interesse pratico e teorico del filtro adattato, vediamo altre metodologie di calcolo.

Si voglia massimizzare il SNR con un filtro lineare. Sia questo filtro B definito da una risposta impulsiva (non necessariamente causale) $\{a_i\}$; cioè

$$(9.23) \quad \frac{A \cdot \sum_i w_i \cdot b_i}{\sigma_\varepsilon^2 \cdot \sum_i |b_i|^2}$$

Per la disuguaglianza di Schwartz³⁰, il massimo si ha se $b_i \propto w_i$ (di cui $b_i = w_i$ è un caso particolare).

³⁰ Ricordiamo che la **disuguaglianza di Schwartz** (o di **Cauchy-Schwartz**) è

$$\left| \int_{-\infty}^{\infty} x(t) \cdot y(t) \cdot dt \right|^2 \leq \int_{-\infty}^{\infty} |x(t)|^2 dt \cdot \int_{-\infty}^{\infty} |y(t)|^2 dt$$

e, nel discreto,

$$\left| \sum_i x_i \cdot y_i \right|^2 \leq \sum_i |x_i|^2 \cdot \sum_i |y_i|^2$$

L'uguaglianza si ha se y è proporzionale a x .

Un altro modo di calcolare il filtro adattato, che giunge allo stesso risultato (9.12), è il seguente:

- si voglia stimare, col metodo dei minimi quadrati, l'ampiezza A di un impulso $\{A \cdot w_i\}$.

Il rumore gaussiano bianco aggiunto può essere visto come un errore gaussiano additivo. La stima sarà lineare per la gaussianità (e potremo indicare lo stimatore con il filtro F di risposta impulsiva $\{f_i\}$). L'errore da minimizzare sarà

$$(9.24) \quad P = E \left[\left(A - \sum_i f_i \cdot y_i \right)^2 \right]$$

Poniamo nella (9.24) $f_i = w_i + a_i$. Abbiamo, ricordando la (9.9),

$$(9.25) \quad \begin{aligned} P &= E \left[\left(A - \sum_i (w_i + a_i) \cdot (A \cdot w_i + \varepsilon_i) \right)^2 \right] = \\ &E \left[\left(A - A \cdot \sum_i w_i^2 - A \cdot \sum_i w_i a_i - \sum_i w_i \varepsilon_i - \sum_i \varepsilon_i a_i \right)^2 \right] = \\ &E \left[\left(A \cdot \sum_i w_i a_i + \sum_i w_i \varepsilon_i + \sum_i \varepsilon_i a_i \right)^2 \right] \end{aligned}$$

Ora, se vogliamo minimizzare P, il primo termine in parentesi, che dipende da A che può essere grande a piacere, si deve annullare, quindi deve essere

$$(9.26) \quad \sum_i w_i a_i$$

(ortogonalità di w e a). Quindi, ricordando che le ε_i sono scorrelate e $E[\varepsilon_i^2] = \sigma_\varepsilon^2$,

$$(9.27) \quad \begin{aligned} E \left[\left(\sum_i (w_i + a_i) \varepsilon_i \right)^2 \right] &= \sigma_\varepsilon^2 \cdot \left(\sum_i (w_i^2 + a_i^2) + 2 \cdot \sum_i w_i \cdot a_i \right) = \\ &= \sigma_\varepsilon^2 \cdot \sum_i (w_i^2 + a_i^2) \end{aligned}$$

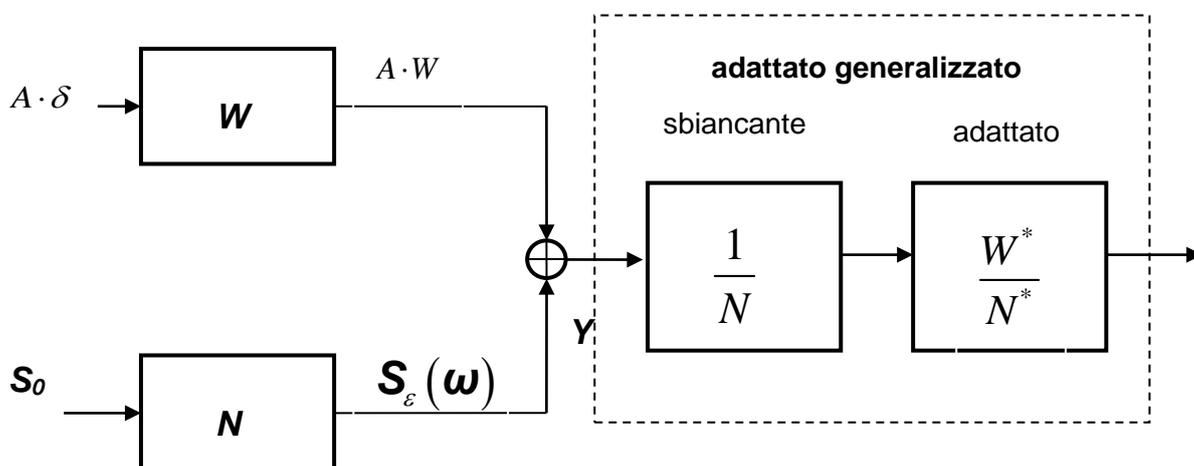
da cui deve essere $a_i = 0$ per tutti gli i e quindi $f_i = w_i$, come avevamo trovato con le altre ipotesi.

Caso generale

Vediamo ora cosa possiamo fare nel caso in cui il rumore non è bianco.

Poiché la nostra soluzione ottima dà un filtro lineare, possiamo “sbiancare” il rumore (e quindi modificare di conseguenza il segnale) e applicare a ciò il metodo appena applicato³¹; attenzione però perché il nuovo filtro andrà adattato al segnale modificato dallo sbiancante.

Per chiarire la situazione, mostriamo lo schema delle operazioni:



³¹ Si deve supporre che il filtro sbiancante sia realizzabile (e stabile). Nella nostra ipotesi di rumore gaussiano stazionario, ciò è sempre possibile.

Se lavoriamo con le frequenze fisiche, per semplicità supporremo $S_0 = 1 \text{ Hz}^{-1}$ (l'ampiezza viene così conglobata in N); se usiamo le frequenze normalizzate (cioè quando usiamo la variabile pulsazione normalizzata Ω), possiamo porre $S_0 = 1$.

In formule, nel dominio della frequenza, abbiamo

$$(9.28) \quad V(\omega) = Y(\omega) \cdot \frac{W^*(\omega)}{|N(\omega)|^2} = Y(\omega) \cdot \frac{W^*(\omega)}{S_\varepsilon(\omega)}$$

Si noti che il filtro “sbiancante” (whitening filter) N è applicato due volte, una volta in modo “causale” (dal passato al futuro) e una volta in modo anti-causale (dal futuro al passato, quando il filtro applicato nel dominio della frequenza è N^*): ciò rende nullo qualsiasi sfasamento che N da solo possa introdurre.

La risposta al segnale, nel dominio della frequenza, è data da

$$(9.29) \quad V(\omega) = \frac{|W(\omega)|^2}{S_\varepsilon(\omega)}$$

Questa è sempre reale positiva e quindi la sua anti-trasformata è una funzione hermitiana.

Se il segnale e il rumore sono reali, è anche una funzione pari.

Se invece c'è solo rumore, all'uscita abbiamo un rumore che ha spettro

$$(9.30) \quad S_v(\omega) = S_0 \cdot \left| \frac{W^*(\omega)}{N^*(\omega)} \right|^2 = S_0 \cdot \frac{|W(\omega)|^2}{S_\varepsilon(\omega)}$$

Si noti che, S_0 è essenziale per motivi di dimensioni fisiche, se usiamo le misure fisiche.

Per calcolare la risposta impulsiva nel tempo e l'autocorrelazione del rumore occorre calcolare

$$(9.31) \quad R_k = \mathcal{F}^{-1} \left(\frac{|W(\omega)|^2}{S_\varepsilon(\omega)} \right)$$

In particolare è importante il valore R_0 , dove ha il massimo assoluto.

Un importante risultato è che la risposta impulsiva in frequenza (9.29) è proporzionale allo spettro di potenza del rumore (9.30), e la risposta impulsiva (9.31) nel tempo è proporzionale all'autocorrelazione del rumore.

L'SNR dopo il filtro è

$$(9.32) \quad SNR = \frac{R_0}{\sqrt{S_0 \cdot R_0}} = \sqrt{\frac{R_0}{S_0}}$$

Teoria della rivelazione (Detection Theory)

Abbiamo costruito il filtro adattato, ma questo dà un valore ad ogni istante di tempo, quindi “mente” quasi sempre: infatti la sua stima è ragionevole solo quando il segnale è effettivamente presente.

Per poter “rivelare” (to detect in Inglese) effettivamente segnali, dobbiamo **decidere** quando il segnale è plausibilmente presente e quando invece no.

Nel caso del problema che abbiamo posto nel paragrafo 0, la decisione viene presa sulla base dell’ampiezza dell’uscita del filtro adattato:

- **decidiamo che ci sia un segnale del tipo cercato se la $\{v_i\}$ ha un massimo relativo e supera una data soglia.**

La situazione è analoga a quella che si incontra nei test statistici.

Supponiamo che l’uscita del filtro abbia una densità di probabilità $f(x; s)$, dove s è l’ampiezza del segnale eventualmente presente. Se poniamo una soglia θ , abbiamo i due casi

A. Il segnale non è presente. Con probabilità

$$(9.33) \quad P_0 = \int_{-\infty}^{\theta} f(x; 0) \cdot dx$$

il segnale non supererà la soglia (e quindi abbiamo indovinato) e con probabilità $1 - P_0$ la soglia viene superata anche in assenza di segnale e quindi commettiamo un **errore di primo tipo**, ovvero abbiamo un “**falso allarme**” (false alarm).

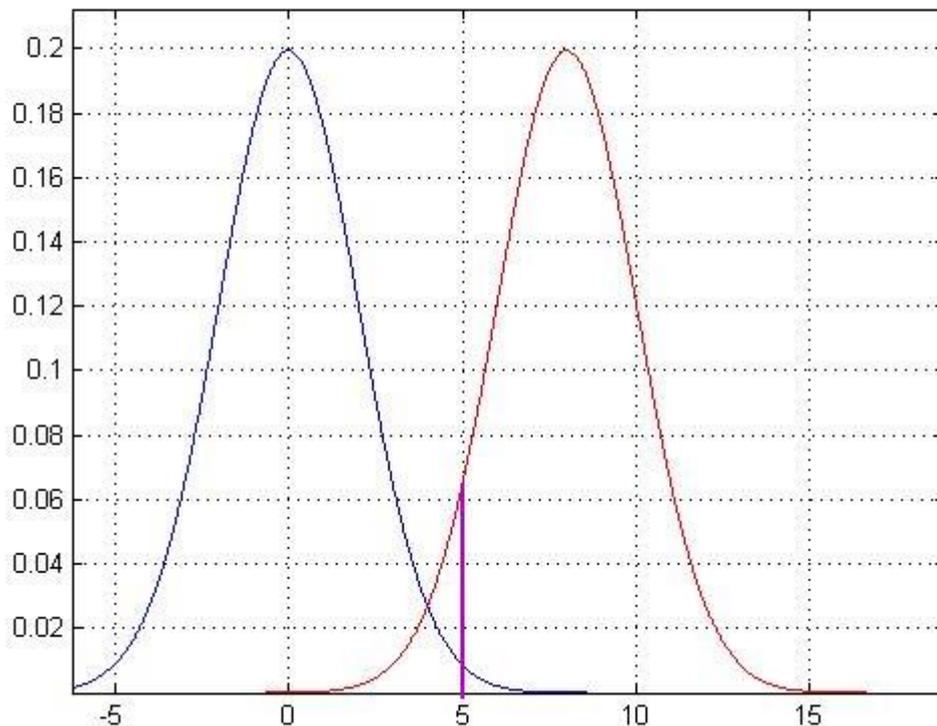
B. Il segnale è presente. Con probabilità

$$(9.34) \quad P_s = \int_{\theta}^{\infty} f(x; s) \cdot dx$$

il segnale supererà la soglia (e quindi abbiamo indovinato) e con probabilità $1 - P_s$ la soglia non viene superata anche in presenza di segnale che quindi non viene rivelato; in tal caso commettiamo un **errore di secondo tipo**, ovvero abbiamo un **“falso rigetto”** (false dismissal).

Se “alziamo” la soglia, cioè aumentiamo θ , possiamo ridurre la probabilità di errori del primo tipo, ma aumentando la probabilità di errori del secondo tipo.

Nella figura seguente è presentato un caso



In questo grafico sono rappresentate le distribuzioni dell’uscita di un rivelatore (che potrebbe essere un filtro adattato), nell’ipotesi di assenza di segnale e nell’ipotesi in cui sia presente un segnale di ampiezza 8.

In questo caso il rumore è additivo (vedi in seguito l’equazione **Error! Reference source not found.**) e gaussiano, di deviazione standard 2. Poniamo la soglia a 5 ed abbiamo

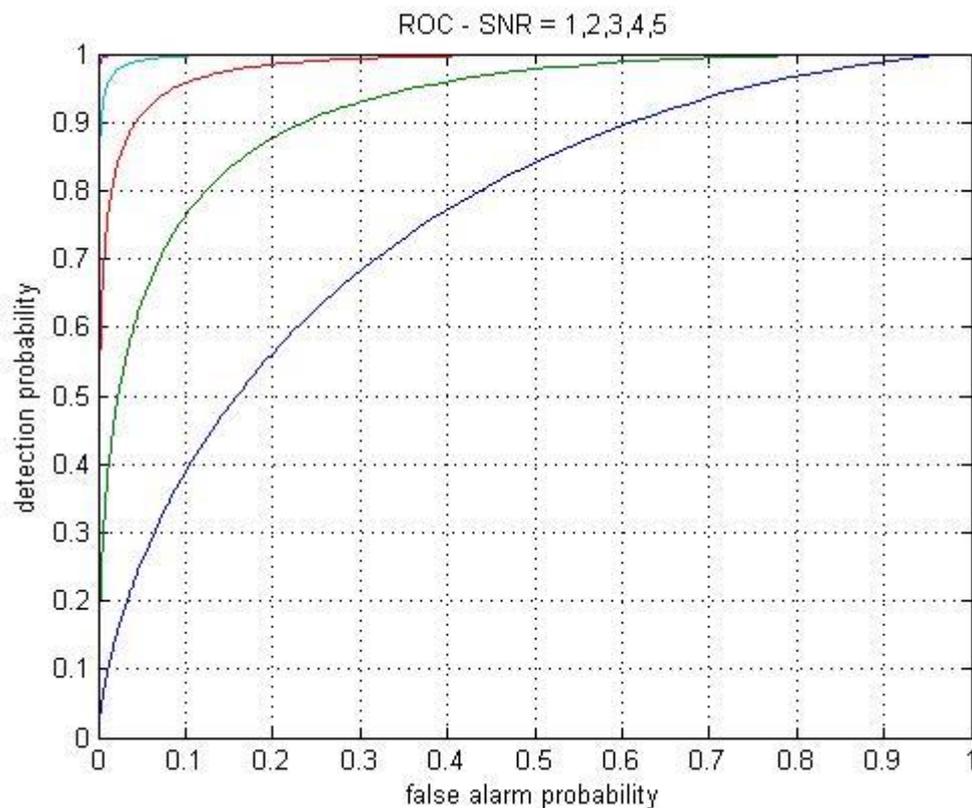
- **probabilità di corretto “rigetto”:** $P_0 = 0.99379$

- probabilità di falso allarme: $1 - P_0 = 0.00621$
- probabilità di corretta rivelazione: $P_1 = 0.933193$
- probabilità di falso rigetto: $1 - P_1 = 0.066807$

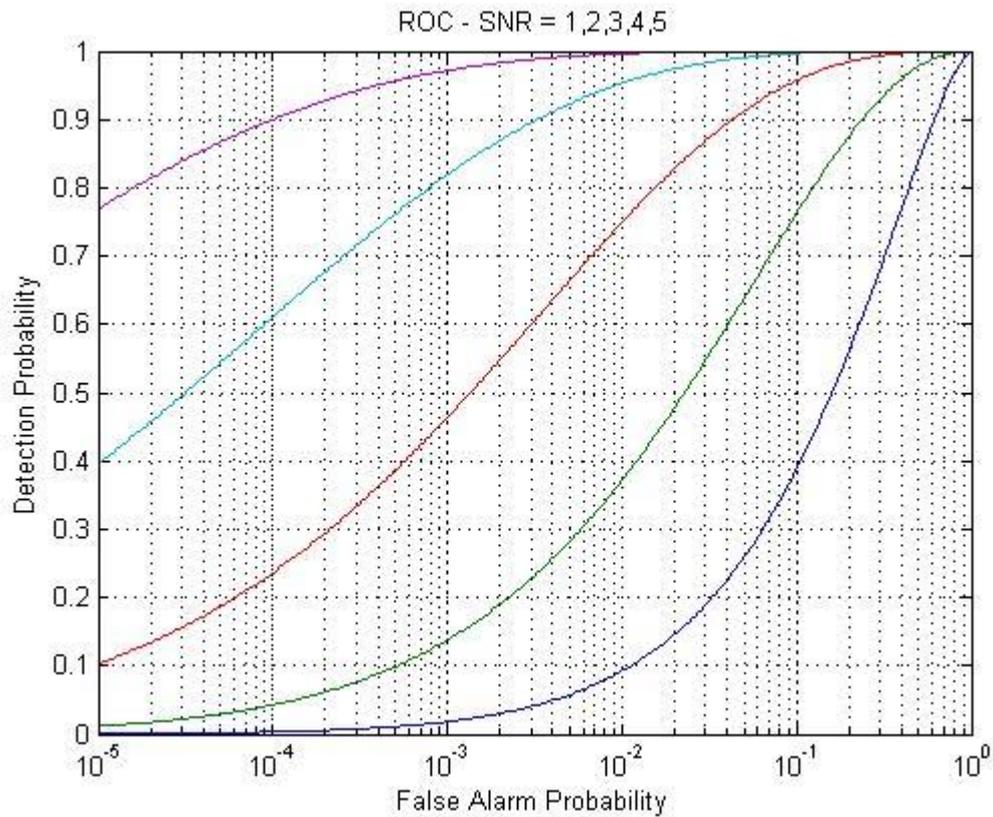
Attenzione: spesso le ipotesi non sono due, ma se il segnale s può assumere molti valori, cioè se s è una variabile casuale (quindi, per esempio, può prendere tutti i valori di un dato intervallo, secondo una data distribuzione di probabilità, che potrebbe anche non essere nota), l'ipotesi A è in effetti molteplice.

Per rappresentare la situazione si usano spesso le **Receiver Operating Characteristics (ROC)**, delle curve che danno, per ogni valore dell'ampiezza del segnale, la probabilità di rivelazione P_s in funzione della probabilità di falso allarme $1 - P_0$, al variare della soglia.

Eccone un esempio, per rumore gaussiano additivo



e, in scala semi-logaritmica,



Se, come accade di frequente, il rumore è additivo, allora

$$(9.35) \quad f(x; s) = f(x - s)$$

e quindi tutto il processo è descritto soltanto da $f(x)$. O, ancora meglio, dalla funzione

$$(9.36) \quad P(x) = \int_x^{\infty} f(\xi) \cdot d\xi = 1 - F(x)$$

che indica la probabilità di falso allarme.

Se vogliamo confrontare due rivelatori (detector) di segnali o due filtri applicati allo stesso rivelatore, basta confrontare le relative due $f(x)$. Se le due $f(x)$ sono dello

stesso tipo, basta confrontarne i parametri. Per esempio, se le $f(x)$ sono gaussiane, basta dare le due deviazioni standard (ed eventualmente i valori aspettati, se ci sono degli offset).

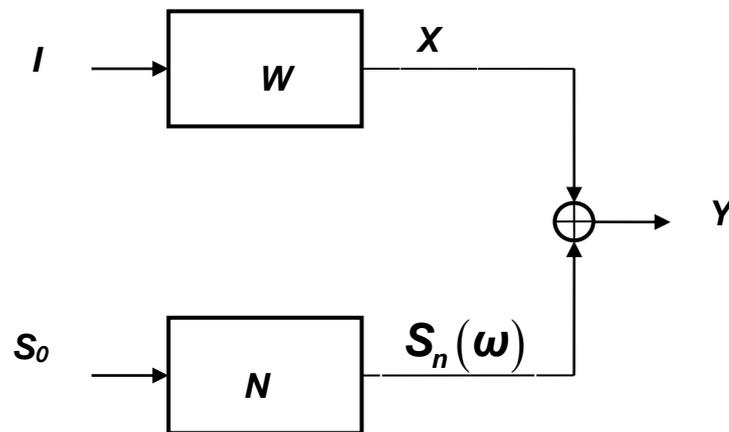
Se lo scopo è il confronto, in tali casi la pletora delle ROC non è una buona scelta.

Filtro di Wiener

Consideriamo ora il problema di stimare un segnale che è modellizzato da un processo stocastico stazionario gaussiano, da dei dati in cui a tale segnale è aggiunto un rumore anch'esso gaussiano e stazionario. Cioè, nel discreto,

$$(9.37) \quad y_i = x_i + n_i$$

Lo schema è il seguente



Il segnale è modellizzato come l'uscita di un filtro lineare W con all'ingresso un segnale con le caratteristiche statistiche di un rumore bianco e che chiamiamo "processo di innovazione".

Dobbiamo costruire la stima ottima nel senso dei minimi quadrati, cioè

$$(9.38) \quad x_i = \sum_{k=-\infty}^{\infty} y_{i-k} \cdot f_k$$

tale che

$$(9.39) \quad E \left[(x_i - \hat{x}_i)^2 \right]$$

sia minimo.

Per il principio di ortogonalità, la condizione di minimo sulla (9.39) porta a che sia

$$E\left[(x_i - \hat{x}_i) \cdot y_k\right] = 0$$

(9.40) ovvero

$$E\left[\left(x_i - \sum_{l=-\infty}^{\infty} (x_{i-l} + n_{i-l}) \cdot f_l\right) \cdot (x_k + n_k)\right] = 0$$

e, passando ai valori aspettati, ricordando la stazionarietà di x e n e che x e n sono indipendenti e quindi scorrelati,

$$(9.41) \quad R_{xx}(k) - \sum_l f_l \cdot (R_{xx}(l+k)) - \sum_l f_l \cdot (R_{nn}(l+k)) = 0$$

Imponendo il minimo errore secondo i minimi quadrati, troviamo, nel dominio della frequenza,

$$(9.42) \quad W(\omega) = \frac{S_x(\omega)}{S_y(\omega)} = \frac{S_x(\omega)}{S_n(\omega) + S_x(\omega)}$$

dove $S_y(\omega)$ è lo spettro di potenza dei dati, $S_x(\omega)$ è lo spettro di potenza del segnale e $S_n(\omega)$ lo spettro di potenza del rumore. Il filtro $W(\omega)$ è detto filtro di Wiener, dal nome del matematico americano Norbert Wiener.

Nel caso più generale, il filtro si scrive

$$(9.43) \quad W(\omega) = \frac{S_{xy}(\omega)}{S_y(\omega)}$$

dove $S_{xy}(\omega)$ è lo spettro incrociato tra il processo x e i dati y.

Realizzazione di filtri

Un filtro, in pratica, non è altro che un sistema discreto. Dei sistemi discreti abbiamo parlato ampiamente. Qui faremo alcune osservazioni sulla loro realizzazione.

Spesso infatti possono esservi varie possibili implementazioni di uno stesso filtro, soprattutto tenendo presente che raramente si possono realizzare filtri “ottimi”, ma ci si deve accontentare approssimazioni, spesso buone. Ciò anche perché l’informazione sui segnali e sui rumori non sempre è completa.

Filtri FIR

I filtri FIR (Finite Impulsive Response) non sono che un altro nome dei sistemi MA (moving average). La caratteristica fondamentale è che questi filtri sono intrinsecamente stabili.

Spesso tuttavia possono essere molto pesanti dal punto di vista computazionale.

Filtri IIR

I filtri IIR (Infinite Impulse Response, AKA filtri AR) sono invece, in genere, meno pesanti computazionalmente (poiché spesso hanno pochi coefficienti) e sono più “naturali” per via della dipendenza a tutte le distanze. In certi casi, piccoli errori nel calcolo dei coefficienti (anche per arrotondamenti) possono dar luogo ad instabilità.

Filtri a sfasamento nullo

I filtri di tipo FIR o IIR sono causali e introducono inevitabilmente (a parte casi banali) uno sfasamento dipendente dalla frequenza.

Talvolta è utile non introdurre sfasamenti, ovvero applicare filtri con funzione di trasferimento reale.

Lo sfasamento introdotto dai filtri FIR può essere lineare con la frequenza, come è lo sfasamento di un semplice ritardo. Infatti la funzione di trasferimento di un ritardo Δt è

$$(9.44) \quad F_D(j\omega; \Delta t) = e^{-j\omega\Delta t}$$

e quindi la fase è

$$(9.45) \quad \varphi = -\omega \cdot \Delta t$$

Ciò accade se i coefficienti del filtro FIR sono simmetrici rispetto a un valore, per esempio i seguenti due:

$$y_i = x_i - 2 \cdot x_{i-1} + x_{i-2}$$

(9.46)

$$y_i = x_i + 2 \cdot x_{i-1} + 2 \cdot x_{i-2} + x_{i-3}$$

Questo ritardo è quasi sempre inessenziale e quindi si parla spesso di filtri a fase lineare come filtri a sfasamento nullo.

Un modo per ottenere filtri a fase nulla, è far passare lo stesso filtro prima in modo causale, cioè dal passato al futuro, e quindi in modo anti-causale, cioè dal futuro al passato. Ovviamente per far ciò dobbiamo conoscere i “dati futuri” (cioè dobbiamo lavorare offline, quindi con un effettivo ritardo). Ciò che accade è che a ciascuna frequenza, se il filtro “causale” la sfasa di φ , quello anti-causale la sfasa di $-\varphi$ e quindi alla fine lo sfasamento è nullo. Ma attenzione! ciascuno dei due filtri applica a quella frequenza un guadagno G , quindi alla fine il guadagno è G^2 .

Vista nel dominio trasformato la situazione è la seguente: se il filtro applicato ha una funzione di trasferimento complessa $F(j\omega)$, applicato in modo anticausale la sua funzione di trasferimento diventa $F^*(j\omega)$ e quindi il filtro effettivamente applicato dopo le due operazioni diventa $|F(j\omega)|^2$.

Filtri in frequenza

Applicare un filtro FIR significa eseguire una convoluzione. Per eseguire una convoluzione di due “vettori”, uno, il pezzo di dati, lungo N e uno, il filtro FIR con M coefficienti, occorre eseguire $M \cdot N$ moltiplicazioni ed altrettante somme. Se M ed N sono numeri abbastanza grandi (per esempio³² $N = 10^9$ ed $M = 10^5$), il carico computazionale diventa elevatissimo.

Si può allora far uso del fatto che nel dominio trasformato di Fourier la convoluzione è una semplice moltiplicazione (quindi, in linea di principio, M volte più veloce).

Dobbiamo però:

- dividere gli N dati in gruppi di lunghezza M (preferibilmente M deve essere una potenza di 2)
- trasformare i dati di ingresso al filtro e quindi anti-trasformare i dati di uscita. E ciò può essere fatto con la FFT (vedi paragrafo 0), con un numero di operazioni proporzionale a $M \cdot \log M$ (invece che M^2 , come sarebbe con la convoluzione).

Con M dell'ordine di 10^5 , si guadagna un fattore circa 1000 nel numero delle operazioni elementari svolte.

Il fatto di lavorare “a pezzi” sui dati, permette di realizzare facilmente filtri non-causali e/o a sfasamento nullo.

Ma attenzione! La FFT, che è un modo veloce di calcolare la DFT, considera i dati periodici di periodo M , e questo creerebbe gravissimi artefatti nel filtraggio. L'escamotage che si usa per questo problema è, in genere, il seguente:

- si divide il pezzo di dati da analizzare in M -pezzi di lunghezza M interallacciati per la metà
- ogni M -pezzo viene utilizzato per analizzarne solo la metà centrale. Infatti con una FFT di lunghezza M , si implementa un filtro FIR di lunghezza effettiva $M/2$ (tipicamente $M/4$ con indice negativo e $M/4$ con indice positivo), e quindi con metà dei coefficienti pari a zero (per filtri non-causali, la parte nulla è la centrale)

³² Sono valori tipici per un giorno di dati di un'antenna gravitazionale, con un filtro ben adattato al rumore reale. Spesso poi si devono applicare contemporaneamente migliaia di filtri.

Filtri adattivi

Nel caso, abbastanza frequente per misure di lunga durata, che le caratteristiche statistiche del rumore cambi in modo non prevedibile, occorre costruire filtri che riconoscano queste variazioni e quindi cambino adattandosi ad esse. Si parla allora di **filtri adattivi** (adaptive filters).

Per costruire un filtro adattato ci sono varie tecniche, dipendentemente dal tipo di variazione del rumore. Per esempio si può avere

- variazione del livello di rumore: lo spettro del rumore rimane della stessa forma, ma cambia la costante moltiplicativa
- variazione della banda del rumore, per esempio se il rumore è un processo del primo ordine, oltre al livello medio, varia il tau dell'autocorrelazione; se è un processo del secondo ordine potrebbe variare la frequenza centrale e/o la larghezza di banda
- variazione complessiva dello spettro, con disturbi a banda stretta che compaiono e scompaiono o variano di caratteristiche.

Nei casi più complessi occorre eseguire delle stime dello spettro di potenza del rumore su tempi adeguati ai tempi di variazione dei disturbi e quindi sulla base di queste costruire il filtro. Talvolta può essere utile costruire più filtri, basati su differenti stime dello spettro di potenza del rumore.

10 - Cenno all'analisi delle immagini

Live Script

Immagini ed elaborazione delle immagini

Un'immagine è descritta da una funzione di due variabili, eventualmente vettoriale se per esempio c'è il colore,

$$(10.1) \quad g(x, y)$$

Distinguiamo tra immagini in bianco e nero e a colori. Nel primo caso in ogni punto la $g(x,y)$ è una funzione scalare, nel secondo a ciascun punto è associato un vettore a 3 dimensioni. Qui ci occuperemo solo di immagini in bianco e nero (anche se esistono modi di visualizzazione a colori di immagini in bianco e nero).

Come avviene per i segnali temporali, così anche le immagini vengono discretizzate per essere "lette" e elaborate dai computer. La funzione $g(x,y)$ diventa quindi una matrice

$$(10.2) \quad g_{ik}$$

Ogni elemento di questa matrice viene detto pixel.

La visualizzazione di una immagine può essere eseguita per esempio per scala di grigi. Nella rappresentazione a scala di grigi, rappresentiamo lo 0 (o il valore minimo) col nero e il valore massimo di nostro interesse col bianco.

Possiamo però anche usare scale "colorate" differenti che talvolta possono risultare più comprensibili o più suggestive.

Eccone un esempio. Qui l'immagine è una foto³³; ogni pixel ha un valore tra 0 e 63.

³³ Non sempre è così: per esempio un'immagine può essere una rappresentazione tempo-frequenza di un segnale.



L'elaborazione delle immagini è un campo molto esteso, in cui si trasforma opportunamente un'immagine, al fine, per esempio, di evidenziare (o nascondere) certe caratteristiche³⁴.

Distinguiamo tre tipi di elaborazione:

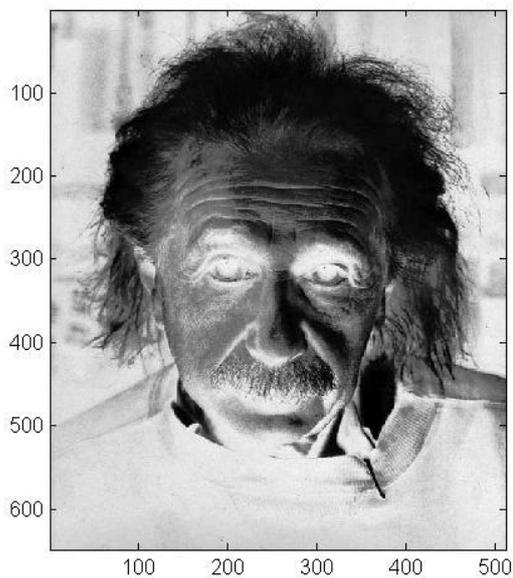
- puntuale : ogni pixel viene modificato dipendentemente dal valore che ha

³⁴ Il fine puo' anche essere artistico.

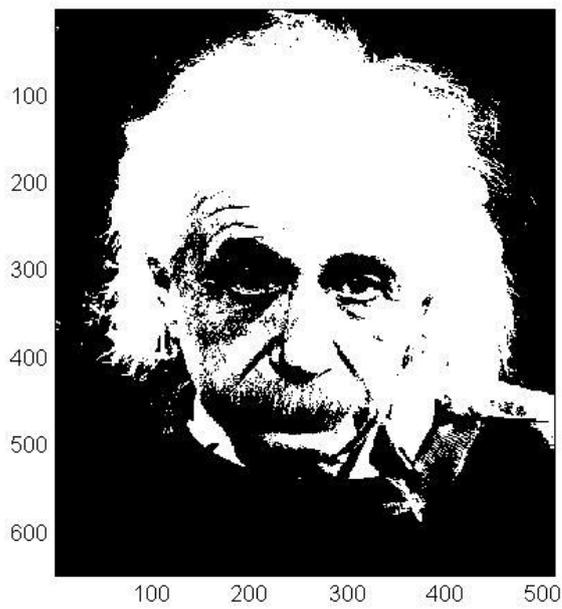
- locale : ogni pixel viene modificato dipendentemente dal valore che hanno quelli vicini (anche non solo gli adiacenti)
- globale : tutti i pixel vengono modificati nello stesso modo (per esempio sommando una costante)

Facciamo un paio di esempi di elaborazione puntuale.

- il negativo : $g'_{ik} = \max(g_{ik}) - g_{ik}$

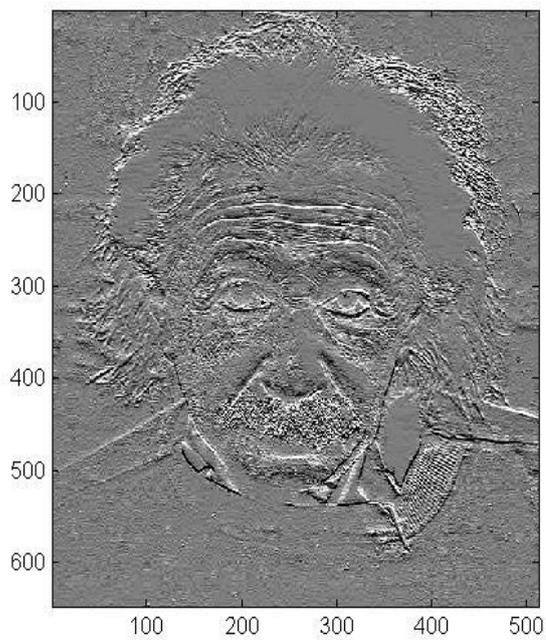


- alto contrasto : $g'_{ik} = \text{sign}(g_{ik} - \theta)$ (dove θ è una soglia)



Vediamo ora una tipica elaborazione “locale”:

- “derivata” : a ogni pixel è sottratto quello che lo precede (per esempio verticalmente)



Elaborazione lineare delle immagini

Le immagini possono essere elaborate in modo analogo ad i segnali. La differenza è che al posto dell'ascissa temporale (o dell'indice "temporale" nel caso dei segnali discreti), abbiamo due ascisse spaziali ortogonali (cioè due indici discreti) (vedi (10.1) e (10.2)).

Analogamente ai segnali temporali, possiamo introdurre l'operazione di convoluzione che nel continuo si definisce

$$(10.3) \quad a(x, y) \otimes b(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(\xi, \zeta) \cdot b(x - \xi, y - \zeta) \cdot d\xi \cdot d\zeta$$

Nel caso discreto si ha

$$(10.4) \quad a_{ik} \otimes b_{ik} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_{im} \cdot b_{(i-m)(k-n)}$$

Elaborazioni di una immagine del tipo della convoluzione vengono dette elaborazioni lineari.

Analogamente al caso dei segnali temporali, possiamo definire la trasformata di Fourier di una immagine $g(x, y)$ come

$$(10.5) \quad G(u, v) = \mathcal{F}(g(x, y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot e^{-j(u \cdot x + v \cdot y)} \cdot dx \cdot dy$$

(si suppone $g(x, y) = 0$ dove non esiste l'immagine). La sua inversa è

$$(10.6) \quad g(x, y) = \mathcal{F}^{-1}(G(u, v)) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(u, v) \cdot e^{j(u \cdot x + v \cdot y)} \cdot dx \cdot dy$$

e, nel caso discreto, equivalentemente alla DFT,

$$(10.7) \quad G(\Phi, \Psi) = \sum_{i=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g_{ik} \cdot e^{-j(i \cdot \Phi + k \cdot \Psi)}$$

e la sua inversa

$$(10.8) \quad g_{ik} = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\Phi, \Psi) \cdot e^{j(i\Phi+k\Psi)} \cdot d\Phi \cdot d\Psi$$

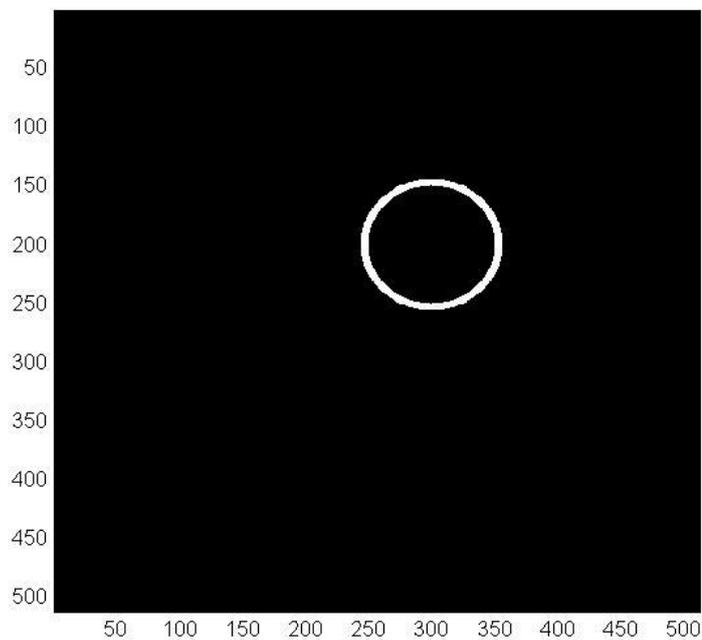
Possiamo definire l'analogo dello spettro di potenza come

$$(10.9) \quad S(u, v) = k \cdot |G(u, v)|^2$$

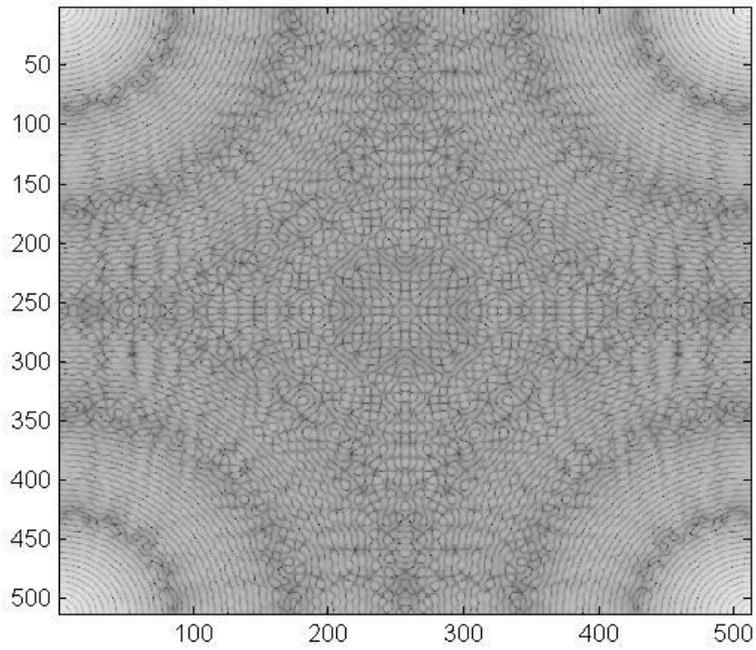
dove k è una costante; chiamiamolo semplicemente spettro, perché ovviamente non è uno spettro di potenza (né di energia). Per l'autocorrelazione abbiamo

$$(10.10) \quad R(\xi, \zeta) = \mathcal{F}^{-1}(S(u, v))$$

Facciamo un esempio. Ecco l'immagine di un cerchio

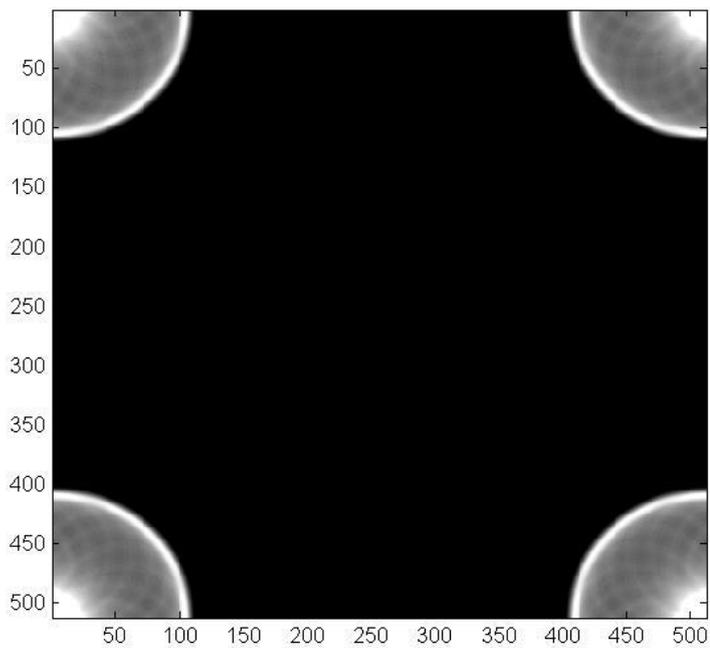


Ed ecco il logaritmo dello spettro



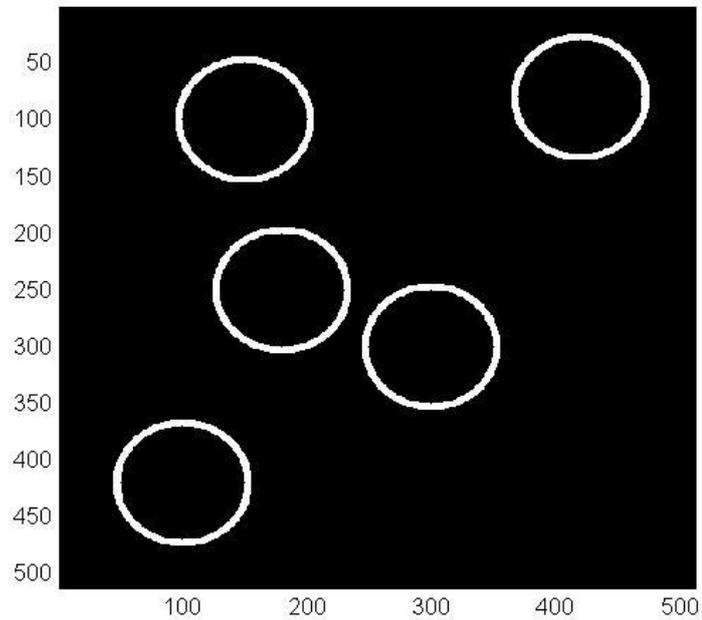
Si noti che questo “spettro” è indipendente dalla posizione del cerchio nella figura (anzi, sul “toro” costruito unendo i due lati verticali e quindi i due orizzontali).

Ed ecco l’autocorrelazione:



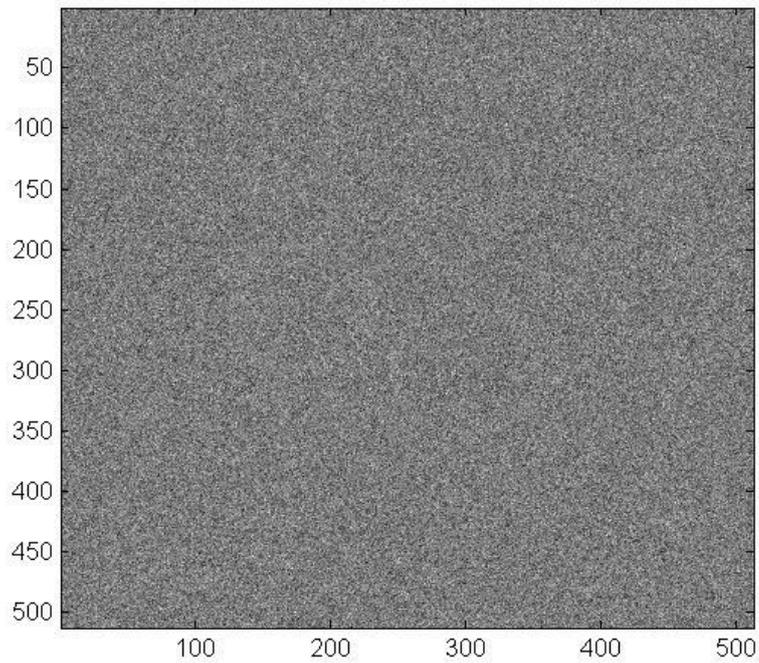
Vediamo ora come migliorare il rapporto segnale rumore in una immagine.

Costruiamo la seguente immagine



Sono 5 cerchi uguali a quello precedente di cui abbiamo fatto lo spettro.

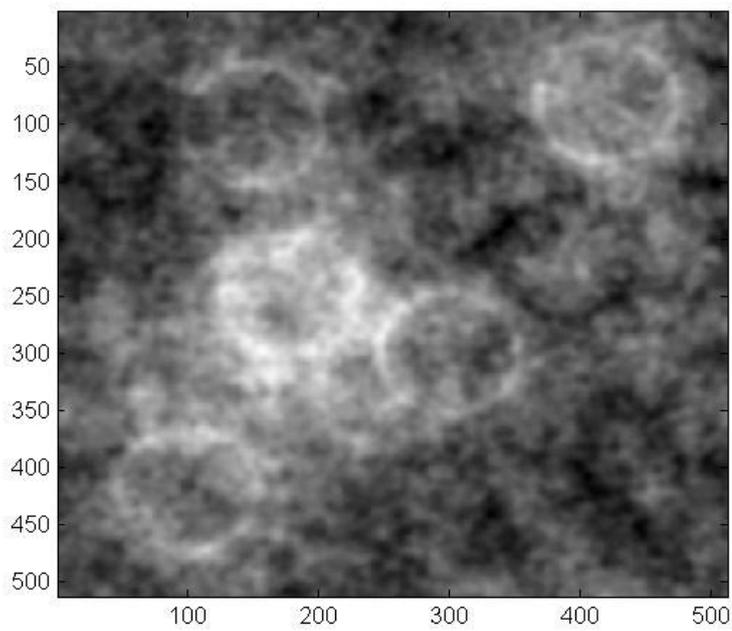
Ora sommiamo a questa immagine del “rumore”, cioè a ciascun pixel una variabile normale. Se il rumore aggiunto è abbastanza “grande”, i cerchi scompariranno. Ecco il risultato avendo aggiunto rumore gaussiano con deviazione standard 4 volte maggiore:



Ora costruiamo un semplice filtro così fatto:

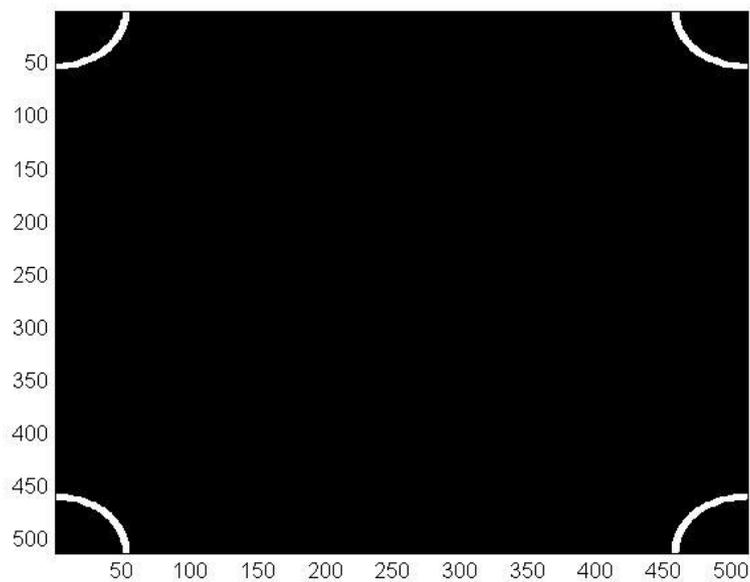
- si prende la trasformata di Fourier dell'immagine disturbata
- si moltiplica per lo spettro del singolo cerchio (posto nel centro "vero" del toro, cioè nel punto $(0,0)$ adiacente agli altri tre spigoli)
- se ne fa la trasformata inversa

Ed ecco il risultato:



Come si vede, anche se disturbati, i cerchi sono ben visibili, nella posizione giusta.

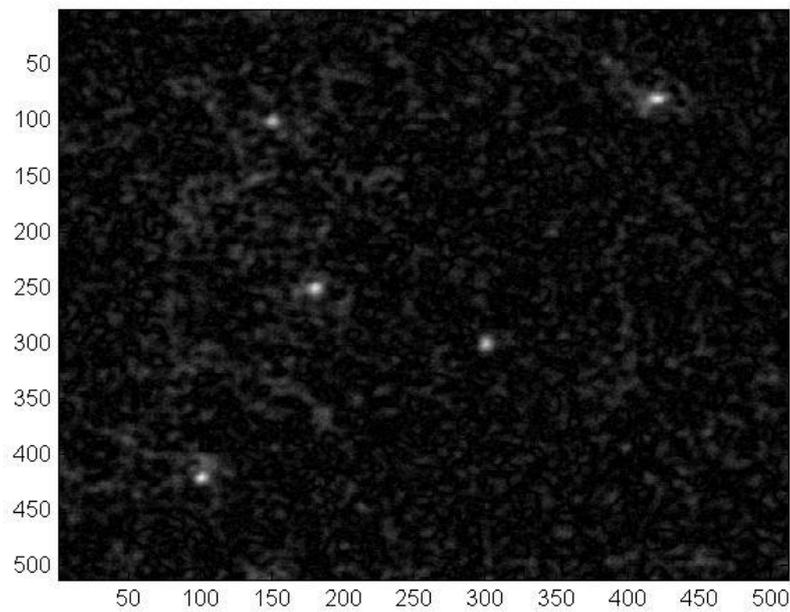
Questo è un filtro piuttosto rozzo. Possiamo fare il filtro adattato (del tutto analogo a quello nel caso dei segnali), prendendo un cerchio nell'origine



e quindi usarlo come “filtro adattato” ed applicarlo all’immagine, eventualmente lavorando con le trasformate di Fourier, cioè essendo M questa immagine e X l’immagine originale, otteniamo

$$(10.11) \quad Y = \mathcal{F}^{-1} \left[\mathcal{F}(X) \cdot \mathcal{F}(M)^* \right]$$

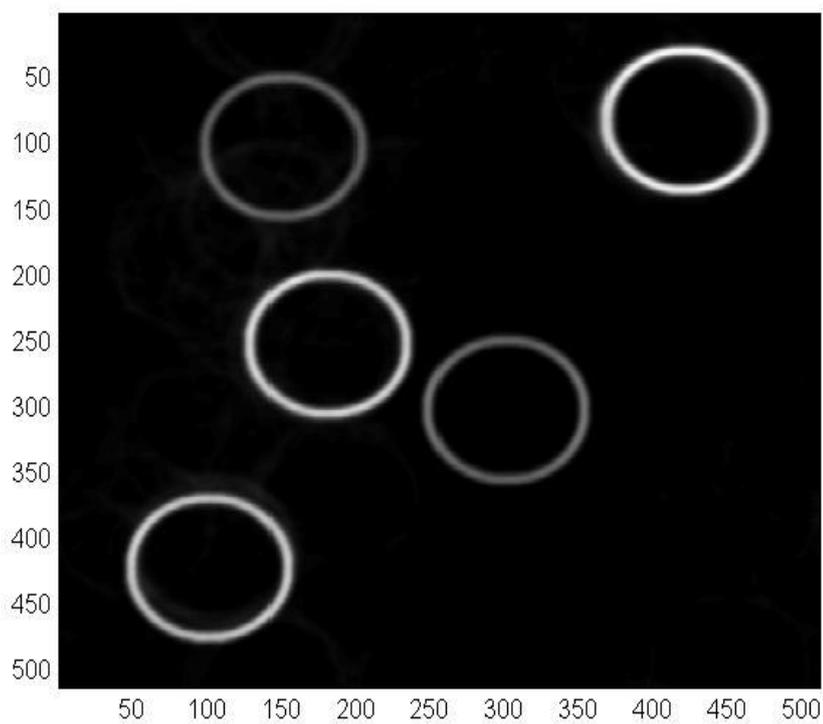
Otteniamo



dove, come si vede, appaiono dei punti luminosi nei centri dei cerchi. Da questa, se il rapporto segnale-rumore è abbastanza buono, possiamo riottenere l’immagine originale con

$$(10.12) \quad X' = \mathcal{F}^{-1} \left[\mathcal{F}(Y \cdot |Y|^3) \cdot \mathcal{F}(M) \right]$$

ottenendo



che è decisamente un buon risultato. Notare che quest'ultima operazione è non-lineare, e, come spesso accade con i metodi non lineari, dà buoni risultati solo se il rapporto segnale-rumore è abbastanza elevato. Se si operasse in modo lineare, se cioè si usasse

$$(10.13) \quad X' = \mathcal{F}^{-1}[\mathcal{F}(Y) \cdot \mathcal{F}(M)]$$

avremmo il risultato originale.

La compressione JPEG

11 - Cenno alla Teoria dell'Informazione

Live Script

Schema di trasmissione di un segnale

Entropia e informazione

Compressione

Codici a correzione

12 - Introduzione all'apprendimento automatico (Machine Learning)

Live Script

Appendice - Trasformate z

Proprietà

	Dominio del tempo	Dominio Z	Dimostrazione	ROC
Notazione	$x[n] = \mathcal{Z}^{-1}\{X(z)\}$	$X(z) = \mathcal{Z}\{x[n]\}$		ROC: $r_2 < z < r_1$
Linearità	$a_1 x_1[n] + a_2 x_2[n]$	$a_1 X_1(z) + a_2 X_2(z)$	$X(z) = \sum_{n=-\infty}^{\infty} (a_1 x_1(n) + a_2 x_2(n)) z^{-n}$ $= a_1 \sum_{n=-\infty}^{\infty} x_1(n) z^{-n} + a_2 \sum_{n=-\infty}^{\infty} x_2(n) z^{-n}$ $= a_1 X_1(z) + a_2 X_2(z)$	Almeno la regione di intersezione di ROC1 e ROC2
Espansione temporale	$x^{(k)}[n] = \begin{cases} x[r], & n = rk \\ 0, & n \neq rk \end{cases}$ <p>r intero</p>	$X(z^k)$	$X_k(z) = \sum_{n=-\infty}^{\infty} x^{(k)}(n) z^{-n}$ $= \sum_{r=-\infty}^{\infty} x(r) z^{-rk}$ $= \sum_{r=-\infty}^{\infty} x(r) (z^k)^{-r}$ $= X(z^k)$	$r^{1/k}$
Traslazione temporale	$x[n - k]$	$z^{-k} X(z)$	$Z\{x[n - k]\} = \sum_{n=0}^{\infty} x[n - k] z^{-n}$ <p>Posto $j = n - k$</p> $\sum_{n=0}^{\infty} x[n - k] z^{-n} = z^{-k} \sum_{j=-k}^{\infty} x[j] z^{-j}$ <p>essendo $x[\beta] = 0$ per $\beta < 0$</p> $z^{-k} \sum_{j=0}^{\infty} x[j] z^{-j} = \sum_{n=0}^{\infty} x[n - k] z^{-n}$ <p>si ha:</p> $= \sum_{j=-k}^{\infty} x[j] z^{-(j+k)} = \sum_{j=-k}^{\infty} x[j] z^{-j} z^{-k}$ $z^{-j} = z^{-k} \sum_{j=0}^{\infty} x[j] z^{-j}$ <p>0 se $\beta < 0$. Da cui:</p> $= z^{-k} X(z)$	ROC, eccetto $\alpha = 0$ se $k > 0$ e $z = \infty$ e $k < 0$ se
Scalatura nel dominio z	$a^n x[n]$	$X(a^{-1}z)$	$Z\{a^n x[n]\} = \sum_{n=-\infty}^{\infty} a^n x(n) z^{-n}$ $= \sum_{n=-\infty}^{\infty} x(n) (a^{-1}z)^{-n}$ $= X(a^{-1}z)$	$ a r_2 < z < a r_1$

Inversione temporale	$x[-n]$	$X(z^{-1})$	$\begin{aligned} \mathcal{Z}\{x(-n)\} &= \sum_{n=-\infty}^{\infty} x(-n)z^{-n} \\ &= \sum_{m=-\infty}^{\infty} x(m)z^m \\ &= \sum_{m=-\infty}^{\infty} x(m)(z^{-1})^{-m} \\ &= X(z^{-1}) \end{aligned}$	$\frac{1}{r_1} < z < \frac{1}{r_2}$
Coniugazione complessa	$x^*[n]$	$X^*(z^*)$	$\begin{aligned} \mathcal{Z}\{x^*(n)\} &= \sum_{n=-\infty}^{\infty} x^*(n)z^{-n} \\ &= \sum_{n=-\infty}^{\infty} [x(n)(z^*)^{-n}]^* \\ &= [\sum_{n=-\infty}^{\infty} x(n)(z^*)^{-n}]^* \\ &= X^*(z^*) \end{aligned}$	ROC
Parte reale	$\text{Re}\{x[n]\}$	$\frac{1}{2} [X(z) + X^*(z^*)]$		ROC
Parte immaginaria	$\text{Im}\{x[n]\}$	$\frac{1}{2j} [X(z) - X^*(z^*)]$		ROC
Differenz.	$nx[n]$	$-z \frac{dX(z)}{dz}$	$\begin{aligned} \mathcal{Z}\{nx(n)\} &= \sum_{n=-\infty}^{\infty} nx(n)z^{-n} \\ &= z \sum_{n=-\infty}^{\infty} nx(n)z^{-n-1} \\ &= -z \sum_{n=-\infty}^{\infty} x(n)(-nz^{-n-1}) \\ &= -z \sum_{n=-\infty}^{\infty} x(n) \frac{d}{dz} (z^{-n}) \\ &= -z \frac{dX(z)}{dz} \end{aligned}$	ROC
Convolut.	$x_1[n] * x_2[n]$	$X_1(z)X_2(z)$	$\begin{aligned} \mathcal{Z}\{x_1(n) * x_2(n)\} &= \mathcal{Z}\{\sum_{l=-\infty}^{\infty} x_1(l)x_2(n-l)\} \\ &= \sum_{n=-\infty}^{\infty} [\sum_{l=-\infty}^{\infty} x_1(l)x_2(n-l)]z^{-n} \\ &= \sum_{l=-\infty}^{\infty} x_1(l) \sum_{n=-\infty}^{\infty} x_2(n-l)z^{-n} \\ &= [\sum_{l=-\infty}^{\infty} x_1(l)z^{-l}] [\sum_{n=-\infty}^{\infty} x_2(n)z^{-n}] \\ &= X_1(z)X_2(z) \end{aligned}$	Almeno la regione di intersezione di ROC1 e ROC2
Crosscorr.	$r_{x_1, x_2} = x_1^*[-n] * x_2[n]$	$R_{x_1, x_2}(z) = X_1^*(1/z^*)X_2(z)$		Almeno la regione di intersezione di ROC of $X_1(1/z^*)$ e $X_2(z)$
Prima differenza	$x[n] - x[n-1]$	$(1 - z^{-1})X(z)$		Almeno la regione di intersezione di ROC of $X_1(z)$ e $ z > 0$
Accumulaz.	$\sum_{k=-\infty}^n x[k]$	$\frac{1}{1 - z^{-1}} X(z)$	$\begin{aligned} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^n x[k] z^{-n} &= \sum_{n=-\infty}^{\infty} (x[n] + x[n-1] + \dots + x[-\infty]) z^{-n} \\ &= X(z) \sum_{j=0}^{\infty} z^{-j} \\ &= X(z) \frac{1}{1 - z^{-1}} \end{aligned}$	$k \cdot z^{-n} + z^{-n} + z^{-n-1} + z^{-n-2} + z^{-n-3} \dots$
Moltiplicaz.	$x_1[n]x_2[n]$	$\frac{1}{j2\pi} \oint_C X_1(v)X_2(\frac{z}{v})v^{-1}dv$		-

Teorema di Parseval	$\sum_{n=-\infty}^{\infty} x_1[n]x_2^*[n]$	$\frac{1}{j2\pi} \oint_C X_1(v)X_2^*\left(\frac{1}{v^*}\right)v^{-1}dv$	
---------------------	--	---	--

Trasformata di alcune funzioni notevoli

Siano:

$$\begin{aligned} \blacksquare u[n] &= \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} \\ \blacksquare \delta[n] &= \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \end{aligned}$$

Funzione, $x[n]$	Trasformata Z, $X(z)$	ROC
$\delta[n]$	1	ogni z
$\delta[n - n_0]$	z^{-n_0}	$z \neq 0$
$u[n]$	$\frac{1}{1 - z^{-1}}$	$ z > 1$
$e^{-\alpha n} u[n]$	$\frac{1}{1 - e^{-\alpha} z^{-1}}$	$ z > e^{-\alpha} $
$-u[-n - 1]$	$\frac{1}{1 - z^{-1}}$	$ z < 1$
$nu[n]$	$\frac{z^{-1}}{(1 - z^{-1})^2}$	$ z > 1$
$-nu[-n - 1]$	$\frac{z^{-1}}{(1 - z^{-1})^2}$	$ z < 1$
$n^2 u[n]$	$\frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3}$	$ z > 1$
$-n^2 u[-n - 1]$	$\frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3}$	$ z < 1$
$n^3 u[n]$	$\frac{z^{-1}(1 + 4z^{-1} + z^{-2})}{(1 - z^{-1})^4}$	$ z > 1$
$-n^3 u[-n - 1]$	$\frac{z^{-1}(1 + 4z^{-1} + z^{-2})}{(1 - z^{-1})^4}$	$ z < 1$
$a^n u[n]$	$\frac{1}{1 - az^{-1}}$	$ z > a $
$-a^n u[-n - 1]$	$\frac{1}{1 - az^{-1}}$	$ z < a $
$na^n u[n]$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$ z > a $
$-na^n u[-n - 1]$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$ z < a $
$n^2 a^n u[n]$	$\frac{az^{-1}(1 + az^{-1})}{(1 - az^{-1})^3}$	$ z > a $

$-n^2 a^n u[-n-1]$	$\frac{az^{-1}(1+az^{-1})}{(1-az^{-1})^3}$	$ z < a $
$\cos(\omega_0 n)u[n]$	$\frac{1-z\cos(\omega_0)}{1-2z\cos(\omega_0)+z^2}$	$ z > 1$
$\sin(\omega_0 n)u[n]$	$\frac{z\sin(\omega_0)}{1-2z\cos(\omega_0)+z^2}$	$ z > 1$
$a^n \cos(\omega_0 n)u[n]$	$\frac{1-az\cos(\omega_0)}{1-2az\cos(\omega_0)+a^2z^2}$	$ z > a $
$a^n \sin(\omega_0 n)u[n]$	$\frac{az\sin(\omega_0)}{1-2az\cos(\omega_0)+a^2z^2}$	$ z > a $