Teoria dei Segnali 2

G.V. Pallottino 2009

Dipartimento di Fisica

Università di Roma Sapienza

La precisione di qualsiasi misura fisica trova un limite negli effetti degli errori casuali, cioè del rumore. Nelle misure di grandezze statiche o lentamente variabili intervengono anche gli effetti di derive, in cui intervengono le componenti spettrali di bassissima frequenza del rumore.

Un caso estremo, pure importantissimo, è quello in cui il rumore è tale da rendere difficile rivelare addirittura la manifestazione di un dato fenomeno.

Sicchè lo studio di fattibilità di qualsiasi esperimento fisico richiede preliminarmente la valutazione del rumore associato alle misure, più precisamente del rapporto segnale/rumore che è possibile ottenere. Sebbene ovvio, questo criterio non sempre viene seguito!

Il rumore limita anche le prestazioni di qualsiasi sistema di comunicazioni (nei sistemi telefonici e radio il fenomeno si manifesta nella banda udibile con un caratteristico soffio, dal quale deriva appunto la denominazione di “rumore”).

Risulta dunque evidente l’importanza dei metodi, numerosissimi e assai vari, che nel corso del tempo sono stati sviluppati per migliorare il rapporto segnale/rumore ottenibile cioè per estrarre convenientemente il *segnale utile* dal *rumore* e dai disturbi. Metodi ad hoc per situazioni specifiche e metodi generali di più vasta applicazione.

Cosa sia da considerarsi come segnale e cosa come rumore dipende esclusivamente dal punto di vista degli obiettivi dello sperimentatore.
In una misura di termometria di rumore, per esempio, si considererà *segnale* il rumore di un resistore e *rumore* la sinusoide, costituita da un residuo di rete a 50 Hz, che disturba la misura.

Resta inteso, in ogni caso, che il segnale e il rumore, perchè risultino in qualche modo distinguibili fra loro, debbano presentare qualche proprietà che valga a differenziarli, nel dominio del tempo o della frequenza.

I vari metodi studiati per estrarre un segnale dal rumore differiscono fortemente a seconda della natura e delle caratteristiche di cosa s'intende per segnale e cosa per rumore.

Ma sono tutti in qualche riconducibili all'uno o all'altro
di due punti di vista fondamentali:

- **FILTRAGGIO**
- **CORRELAZIONE**

I metodi di filtraggio mirano generalmente a ricostruire la forma originale di un segnale dopo che è stato corrotto dal rumore.
Per consentire poi la misura di determinati parametri (ampiezza, ecc.) del segnale.
Portando quindi a entrare nel campo della teoria statistica della stima.

I metodi a correlazione mirano piuttosto a rivelare la presenza di un segnale (di forma data) in presenza di rumore (stabilendone, per esempio, il tempo di occorrenza), ma senza necessariamente ricostruirne la forma originale.
Portando quindi a entrare nel campo della teoria statistica della rivelazione.

Ma vedremo anche che in vari casi specifici i due approcci risultano equivalenti.

Il problema che si pone in generale è il seguente:
esprimendo il risultato x(t) di una osservazione sperimentale nella forma

$$x(t) = s(t) + n(t)$$

vogliamo ricostruire il segnale s(t), attraverso una stima $\hat{s}(t)$ basata sull'osservazione x(t), oppure vogliamo stabilire la presenza o meno del segnale stesso (eventualmente determinandone l'ampiezza, il tempo di occorrenza, o altre proprietà).

Si noti che qui abbiamo fatto un'ipotesi importante, non sempre verificata: che il rumore agisca sommandosi al segnale (*rumore additivo*) e non attraverso operazioni nonlineari (per esempio, di moltiplicazione fra segnale e rumore).
Come accadrebbe, per esempio, nel passaggio attraverso un sistema nonlineare di una combinazione lineare di segnale e di rumore.

Esempio di *rumore moltiplicativo*, anzichè additivo, è quanto si verifica nel *fading*, cioè nella attenuazione variabile nel tempo a cui sono soggetti i segnali radio in presenza di propagazione irregolare
(per esempio, variazioni della densità elettronica degli strati riflettenti che assicurano determinate forme di propagazione)
come avviene a volte nell'ascolto di una stazione radio molto lontana.

I metodi sviluppati per attuare quanto detto sopra sono assai vari, differendo sopratutto in relazione alla natura di ciò che consideriamo segnale e cosa consideriamo rumore.

Segnali stazionari oppure transitori?
Rumore a spettro costante o rumore con spettro di forma data?
Un segnale armonico (di frequenza data o incognita?) in presenza di uno spettro di rumore?
Un segnale transitorio di forma nota? (come avviene per es., nel radar)
Segnali in presenza di disturbi impulsivi?
Eccetera.

TECNICHE DI CORRELAZIONE

STIMA DI UNA COSTANTE (FORMA CLASSICA DI FILTRAGGIO)

E' noto dai tempi di Gauss come si procede per stimare una grandezza costante s avendo a disposizione una serie di misure x, che rappresentiamo come una variabile casuale, in modo da minimizzare il valore aspettato dello scarto quadratico.

Si noti che questa non è l'unica scelta possibile,
ma ha il vantaggio che il problema è facilmente trattabile.

Per minimizzare $E[(x - s)^2]$ basta espandere, derivare rispetto ad s e annullare, ottenendo così: $\hat{s} = E[x]$
che ha due proprietà pregevoli:

minimizza lo scarto quadratico medio,

annulla il valore aspettato dello scarto $x-s$

e un evidente difetto: richiede la conoscenza di infiniti valori di x.

Potevamo procedere diversamente, per esempio mirando a minimizzare il valore assoluto dello scarto: in tal caso avremmo ottenuto la mediana di x anzichè la media.

In pratica la stima si può compiere mediando un numero finito n di campioni di x. In tal caso la stima coincide con s in valore aspettato, ma fluttua a sua volta con deviazione standard $\sigma_x^2/\sqrt{n}$
(se i campioni sono indipendenti).

Come si è detto, mediare significa eseguire una media temporale. Cioè idealmente calcolare il valor medio della osservazione x(t) su tempo infinito. In pratica mediando su un tempo finito e dunque ottenendo una stima soggetta a fluttuazioni residue.

In questo caso – tempo di osservazione finito - entrano in gioco le proprietà statistiche dinamiche del rumore (correlazione temporale).

FILTRI DI MEDIA (per ora nel caso di segnale costante)

Cosideriamo l'osservazione come somma di una costante s e di un processo di rumore n(t) a media nulla

$$x(t) = s + n(t)$$

La varianza della stima di s fornita da un filtro di media, a parità di varianza del rumore σ_n^2, dipende in generale sia dal filtro che dalle proprietà statistiche del rumore: spettro e autocorrelazione (che in questo caso coincide con l'autocovarianza).

Se il rumore ha spettro $S_{nn}(\omega)$, autocorrelazione $R_{nn}(\tau)$ e varianza $\sigma_n^2=R_{nn}(0)$, la varianza all'uscita del filtro, cioè la varianza della stima di s, è data in generale dall'espressione:

$$\sigma^2 = R(0) = FT^{-1}[|H(j\omega)|^2 \, S_{nn}(\omega)]$$

dove $H(j\omega)$ è la funzione di trasferimento del filtro.

Il più semplice filtro di media (filtro rettangolare) realizza l'integrazione normalizzata del segnale sul tempo T

$$y(t) = \frac{1}{T} \int_{t-T}^{t} x(t')\,dt'$$

e ha pertanto risposta impulsiva, di forma appunto rettangolare,

$$h(t) = \frac{1}{T}\left(u(t) - u\left(t - T\right)\right)$$

e funzione di trasferimento $H(s) = (1-\exp(-sT))/s$

Spesso si usa anche un filtro con memoria "esponenziale", cioè un filtro del primo ordine con costante di tempo t_o e funzione di trasferimento

$$H(s) = 1/(1 + t_o s)$$

che ha varie denominazioni: filtro RC, filtro a finestra mobile esponenziale, filtro autoregressivo nel caso discreto, ecc.

Un caso interessante (sebbene non molto realistico) è quello di rumore bianco a banda limitata, con spettro (sempre bilatero):

$S_{nn}(\omega) = \sigma_n^2/2B$ per $|\omega| \leq 2\pi B$ $S_{nn}(\omega) = 0$ altrove

autocorrelazione $R_{nn}(\tau) = \sigma_n^{\,2}\left(\frac{\sin(2\pi B\tau)}{2\pi B\tau}\right)$ e varianza σ_n^2

L'autocorrelazione si annulla per qualsiasi $\tau = i/2B$, con i intero >0, sicchè i valori del rumore a distanza 1/2B sono scorrelati fra loro e quindi indipendenti.
Di conseguenza, sommando n campioni equidistanti di x(t) sul tempo di osservazione T = n/2B (oppure integrando sullo stesso intervallo), la varianza residua, cioè la varianza della stima, è semplicemente:

$\sigma^2 = \sigma_n^2/2BT$

Per rumore con spettro arbitrario le cose si complicano.
Si trova (faticosamente) che la varianza del rumore residuo, cioè la varianza della stima, è data dall'espressione:

$$\sigma^2 = \frac{1}{T}\int_{-T}^{T}\left(1-\frac{|\tau|}{T}\right)R_{xx}(\tau)\,d\tau$$

che per T sufficientemente grande si approssima con

$$\sigma^2 \approx \frac{1}{T}\int_{-T}^{T} R_{xx}(\tau)\,d\tau$$

Varianza che si annulla per $T\to\infty$, come era ragionevole aspettarsi.

Si noti che se il rumore ha media nulla, come di solito si assume, allora tutte le autocorrelazioni sono autocovarianze;
altrimenti la stima, è polarizzata (*biased*), cioè affetta da errore sistematico, irriducibile per qualsiasi tempo d'integrazione.

Caso più realistico è quello per cui il rumore è rumore bianco filtrato da un filtro passabasso del primo ordine,

cioè con spettro $S_{nn}(\omega) = 2t_o\sigma_n^2 /(1+\omega^2 t_o^2)$

e autocorrelazione $R_{nn}(\tau) = \sigma_n^2 \exp(-|\tau|/t_o)$

In questo caso la varianza all'uscita del filtro di media rettangolare è:

$$\sigma^2 = \frac{2\sigma_x^2 t_o}{T}\left(1-\frac{t_o}{T}\left(1-\exp\left(\frac{-T}{t_o}\right)\right)\right)$$

che si riduce semplicemente al fattore fuori parentesi quando il tempo d'integrazione T è molto maggiore del tempo di correlazione t_o del rumore.

E allora si ritorna al risultato ottenuto prima per il rumore bianco con banda B (limitata bruscamente),
dato che la banda equivalente del rumore filtrato è $B=1/4t_o$

E se invece di eseguire medie temporali operiamo medie d'insieme?
Questa tecnica utilizza medie eseguite su una molteplicita` di ripetizioni indipendenti di una medesima osservazione sperimentale.
Nella misura, per esempio, della risposta impulsiva di un sistema, quando questa e' soggetta a rumore: eccitando con una sequenza di delta e mediando poi sulle risposte.

O anche nelle misure dei "potenziali evocati" in elettrofisiologia, sempre mediando sulle risposte corrispondenti a una sequenza di eccitazioni.

OSSERVAZIONI SUI FILTRI DI MEDIA

Sono utili anche se il segnale non è una costante, ma varia nel tempo, purchè lentamente rispetto al tempo d'integrazione. In tal caso si parla di filtri di *smoothing* (smussamento), che eliminano dunque le componenti di alta frequenza del rumore, ripulendo il segnale.

Ma allora diventa importante che il filtro di media non introduca ritardo (impossibile ad ottenersi in hardware, con oggetti fisicamente realizzabili, quindi causali, ma ben ottenibile lavorando al calcolatore)

Il filtro a finestra mobile dovrà allora eseguire sull'osservazione la seguente operazione (di media "centrata" e non ritardata):

$$y(t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} x(t')\, dt'$$

I filtri di media vanno in crisi se il rumore non si comporta bene: per esempio in presenza di disturbi impulsivi, occasionali, di grande ampiezza. Sia nella stima di un segnale costante che, sopratutto, di uno variabile.

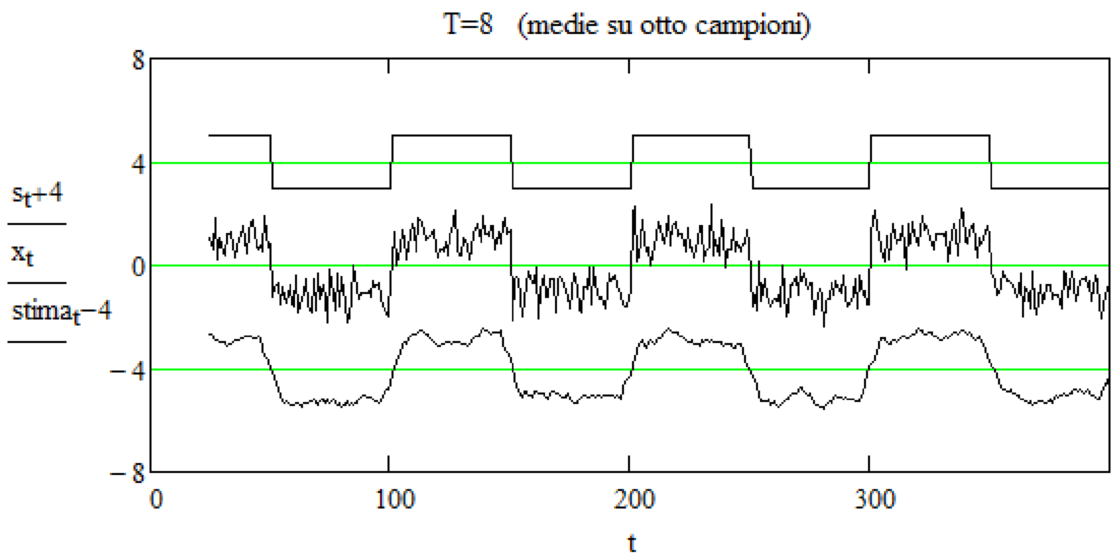
In questo caso può convenire eseguire operazioni nonlineari.
Per esempio con un filtro a mediana, ben realizzabile per segnali a tempo discreto. Cioè un filtro che fornisca, a ogni istante, la mediana degli n+1 campioni dell'osservazione (n/2 prima, quello corrente ed n/2 dopo).

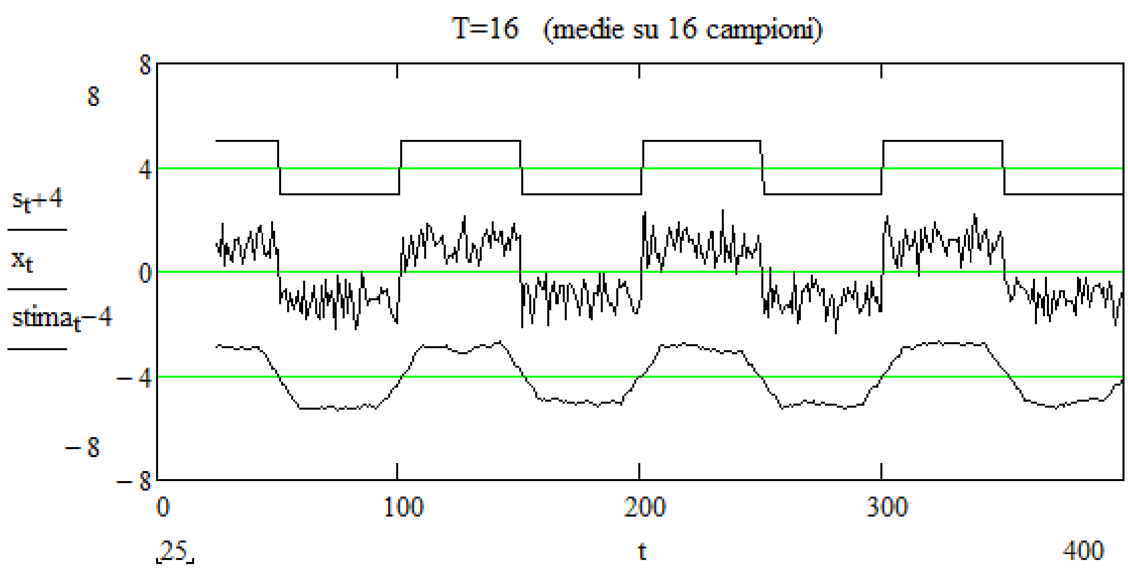
Oppure con un filtro del tipo detto α, che calcola il valor medio dei dati in un intervallo T, escludendo però le code della distribuzione locale (cioè considerando solo i dati nell'intervallo per cui $F(x)>\alpha$ e $F(x)<1-\alpha$)

Ricordiamo del resto che in certi apparati si usa il *noise blanker*:
un dispositivo che annulla semplicemente la trasmissione quando il rumore diventa eccessivamente elevato.

ESEMPIO

filtro di smoothing applicato a un’onda quadra rumorosa: la riduzione del rumore è accompagnata da allungamento dei tempi di transizione

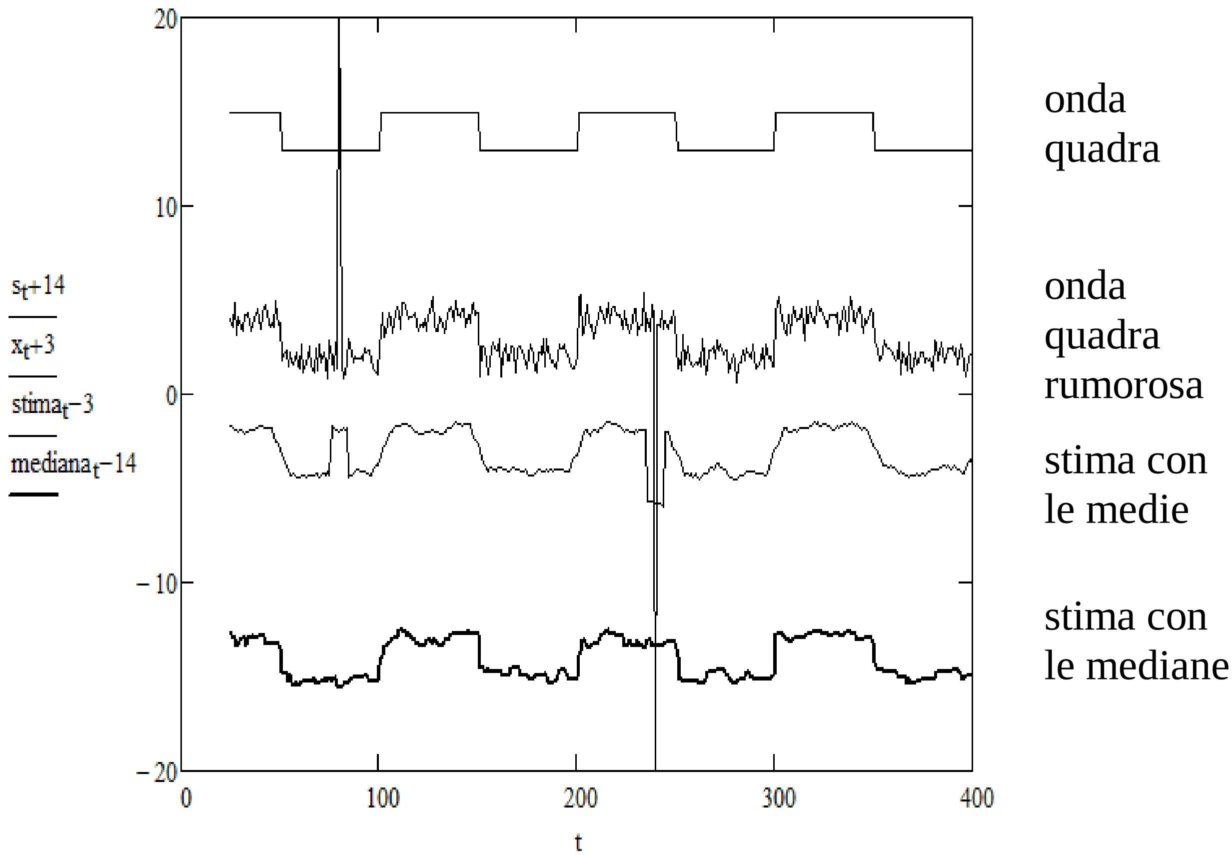




E’ chiaro che al crescere del tempo d’integrazione T, l’errore di stima complessivo prima diminuisce e poi aumenta di nuovo (quando l’effetto dell’integrazione distorce il segnale, trasformando l’onda rettangolare in trapezoidale).

Nel caso di segnali soggetti a rumore impulsivo conviene usare un filtro di smoothing a mediana

Qui sotto è mostrata l'inefficacia di un filtro di media (T = 16 s) e l'efficacia di un filtro a mediana (operante su 7 campioni alla volta) per il segnale a onda quadra rimorosa visto prima,
in presenza di due disturbi impulsivi di grande ampiezza



Osservazione:
il filtraggio nonlineare offre prestazioni ottime in casi particolari
ma si tratta di soluzioni ad hoc,
dal momento che la teoria del filtraggio nonlineare è ben lungi da un livello di sviluppo, sopratutto di generalità,
confrontabile con quella del filtraggio lineare

FILTRAGGIO DI SEGNALI

E’ basato in generale sulla diversa distribuzione in frequenza del segnale e del rumore.
Come ci aspettiamo che sia fatto il filtro ottimo?

Torniamo alla stima di una costante s,
che facciamo usando una combinazione lineare di n variabili casuali x_i che costituiscono i dati di partenza

$$\hat{s} = a_1x_1 + a_2x_2 + \ ... + a_nx_n$$

scegliendo le costanti a_i in modo da minimizzare l’errore quadratico medio di stima $Q = E[(s - \hat{s})^2]$

Principio di ortogonalità
l’errore quadratico medio anzidetto Q è minimo quando le costanti sono tali che l’errore sia ortogonale ai dati:

$$E[((s - \hat{s})x_i] = 0 \qquad \text{per} \quad i = 1, 2,... n$$

come si dimostra annullando le derivate di Q rispetto ad a_i

E veniamo al caso che ci interessa: la stima $\hat{s}(t)$ di un segnale s(t) in base all’osservazione x(t) = s(t) + n(t)

Ora la stima è una funzione del tempo, e i dati disponibili non sono numeri, ma anch'essi una funzione del tempo: x(t)

Qui notiamo che il problema del filtraggio è un caso particolare di un problema di stima più generale,
nel quale, oltre al filtraggio, rientrano:
la *predizione*, stabilire i valori futuri di s(t) (per t>t*) conoscendo x(t) (oppure s(t)) solo fino all’istante t*
l’interpolazione,
ecc.

Esempio particolare
Vogliamo stimare il segnale s(t) in base al valore corrente dell'osservazione $x(t) = s(t) + n(t)$

cioè vogliamo trovare la costante a che ci serve per ottenere
$\hat{s}(t) = ax(t)$ in modo da minimizzare l'errore di stima

Applicando il principio di ortogonalità abbiamo

$E[(s(t) - ax(t))\, x(t)] = 0$ che riscriviamo nella forma

$R_{sx}(0) - aR_{xx}(0) = 0$

da cui si ottiene immediatamente $a = R_{sx}(0)/R_{xx}(0)$

Se il rumore e il segnale non sono correlati (cioè $R_{sn}(\tau)=0$), allora

$R_{sx}(0) = R_{ss}(0)$ (varianza del segnale)
$R_{xx}(0) = R_{ss}(0) + R_{nn}(0)$ (somma delle varianze del segnale e del rumore)
e quindi $a = R_{ss}(0)/(R_{ss}(0) + R_{nn}(0))$

L'errore quadratico medio di stima si ricava nella forma

$Q = E[(s - \hat{s})^2] = R_{ss}(0)R_{nn}(0)/[R_{ss}(0) + R_{nn}(0)]$

Si noti che l'errore di stima Q si può calcolare direttamente dalla sua definizione $Q = E[(s - \hat{s})^2]$,
oppure utilizzando il principio di ortogonalità, che stabilisce l'ortogonalità fra scarto e osservazione, e dunque fra lo scarto e qualsiasi combinazione lineare delle osservazioni, quale è appunto la stima. Così procedendo si ottiene l'importante risultato:

$$Q = E[(s - \hat{s})s]$$

Assai più interessante è il caso della stima dell'andamento del segnale s(t) in un certo intervallo di tempo (cioè la stima di s(t) per ciascun valore di t) conoscendo l'osservazione x(t) in tale intervallo, fra a e b.

In questo caso non avremo una combinazione lineare di variabili casuali, ma un integrale del prodotto dell'osservazione per una opportuna funzione di pesatura:

$$\hat{s}(t) = \int_a^b h(\xi)\, x(\xi)\, d\xi$$

La forma della funzione di pesatura si ottiene imponendo ancora la minimizzazione dell'errore quadratico medio,
applicando il principio di ortogonalità:

$$E[\left(s(t) - \int_a^b h(\alpha)\, x(\alpha)\, d\alpha \right) x(\xi)] = 0 \qquad \text{per} \quad a \le \xi \le b$$

che si traduce nel risultato

$$R_{sx}(t-\xi) = \int_a^b R_{xx}(\alpha-\xi)\, h(\alpha)\, d\alpha$$

(ottenuto nell'ipotesi di stazionarietà dei processi)

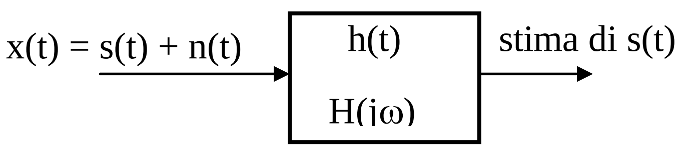
Da questa equazione integrale si può ricavare la funzione di pesatura h
Impresa tutt'altro che banale analiticamente, in generale, ma sempre attuabile numericamente, approssimando l'integrale con una somma.

TEORIA DI WIENER-KOLMOGOROFF

Ammettiamo ora di conoscere l'osservazione x(t) su tutto l'asse dei tempi. In tal caso, sempre nell'ipotesi di stazionarietà dei processi che ci permette di scrivere h(t, ξ) come h(t-ξ), abbiamo

$$\hat{s}(t) = \int_{-\infty}^{\infty} h(t-\xi)\, x(\xi)\, d\xi = \int_{-\infty}^{\infty} x(t-\alpha)\, h(\alpha)\, d\alpha$$

Lo stimatore del processo può essere interpretato come l'uscita di un sistema lineare e stazionario con risposta impulsiva h(t) il cui ingresso è costituito dal segnale osservato x(t).
Si noti che tale sistema non è causale, cioè non è realizzabile fisicamente, dato che a ogni istante la stima richiede la conoscenza del futuro, oltre che del passato (h(t) $\neq$ 0 per t<0).



Applicando ancora il principio di ortogonalità e ponendo τ = t - ξ ricaviamo l'equazione integrale

$$R_{sx}(\tau) = \int_{-\infty}^{\infty} R_{xx}(\tau-\alpha) h(\alpha)\, d\alpha \qquad \text{valida per qualsiasi } \tau$$

La soluzione è immediata trasformando secondo Fourier ambo i membri (si ricordi il teorema di convoluzione):

$S_{sx}(\omega) = S_{xx}(\omega)\, H(j\omega)$

quindi la funzione di trasferimento cercata è $\qquad H(j\omega) = \dfrac{S_{sx}(\omega)}{S_{xx}(\omega)}$

antitrasformando la quale si ricava h(t).

Se il segnale e il rumore hanno il buon gusto di essere indipendenti, come del resto avviene frequentemente, allora $S_{sn}(\omega)=0$ e quindi

$$S_{sx}(\omega) = S_{ss}(\omega) + S_{sn}(\omega) = S_{ss}(\omega)$$
$$S_{xx}(\omega) = S_{ss}(\omega) + S_{sn}(\omega) + S_{ns}(\omega) + S_{nn}(\omega) = S_{ss}(\omega) + S_{nn}(\omega)$$

In tal caso si ottiene

(funzione reale di ω) $$H(j\omega) = \frac{S_{ss}(\omega)}{S_{ss}(\omega) + S_{nn}(\omega)}$$

il cui significato fisico è immediato
e anche, una volta tanto, in ottimo accordo con l'intuizione,
come pure con le tecniche di filtraggio viste precedentemente, usate per stimare una costante o un segnale lentamente variabile.

L'errore di stima? Si calcola usando l'espressione $Q = E[(s - \hat{s})\, s]$ ottenendo:

$$Q = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_{ss}(\omega)\, S_{nn}(\omega)}{S_{ss}(\omega) + S_{nn}(\omega)}\, d\omega$$

Si vede immediatamente che l'errore di stima si annulla quando il segnale e il rumore occupano bande di frequenza diverse, senza sovrapposizione
(e allora $S_{ss}(\omega)\, S_{nn}(\omega) = 0$ per qualsiasi ω).

In questo interessante caso particolare

dove $S_{ss}(\omega) \neq 0$ e $S_{nn}(\omega) = 0$ sarà $H(j\omega) = 1$
dove $S_{ss}(\omega) = 0$ e $S_{nn}(\omega) \neq 0$ sarà $H(j\omega) = 0$

Estrazione di una sinusoide immersa nel rumore:
occorre un filtro con larghezza di banda infinitamente stretta
(Evidenti problemi di realizzazione, ma comunque disponibilità pratica di dati solo su un tempo finito, da cui dipenderà la banda effettiva e l'errore residuo. Lo stesso problema che si pone nell'analisi spettrale).

ESEMPI

Segnale a banda limitata (taglio a ω_s) in presenza di rumore bianco:

$S(\omega) = S/(1+(\omega/\omega_s)^2)$ $N(\omega) = N$

$$H(j\omega) = \frac{S(\omega)}{S(\omega)+N} = \frac{S}{S+N}\,\frac{1}{1+\frac{N}{S+N}\left(\frac{\omega}{\omega_s}\right)^2}$$

Si noti che la pulsazione di taglio del filtro non è ω_s
ma è più alta, tanto più quanto è maggiore il rapporto S/N
(per alti valori di S/N il filtro, giustamente, allarga la banda):

$\beta = \omega_s\,(1+S/N)^{1/2}$

Come è fatto il filtro nel dominio del tempo? Antitrasformando si ottiene la risposta impulsiva, che è non causale, dato che segue la legge esponenziale bilatera

$h(t) = A\exp(-\beta|t|)$ dove β è la pulsazione di taglio

che rappresenta un particolare filtro di media,
che fra l'altro gode della proprietà di non introdurre ritardo

E se volessimo filtrare del rumore bianco in presenza di una sinusoide?

TECNICHE DI CORRELAZIONE

Qui possiamo distinguere due casi fondamentali:

1) quando si vuole individuare la presenza di un segnale
(ed eventualmente stimarne determinati parametri)
disponendo di un *segnale di riferimento*, costituito dalla forma, nota, del segnale stesso

per esempio, sappiamo che il segnale ha forma rettangolare di durata T, ma non ne conosciamo l'ampiezza e il tempo di occorrenza

2) quando si vuole individuare la presenza di un segnale senza disporre di alcun segnale di riferimento, utilizzando allora correlazioni fra osservazioni diverse dello stesso segnale:
- a tempi diversi, ma disponendo di informazioni sul tempo d'inizio di ciascuna ripetizione del segnale
- agli stessi tempi, ma utilizzando due o più diverse sorgenti di segnale

ESTRAZIONE DI UNA SINUSOIDE

Conosciamo la frequenza angolare ω, ma non la fase nè l'ampiezza, di una sinusoide immersa nel rumore. Vogliamo stimarne l'ampiezza.

La soluzione consiste nell'eseguire la correlazione incrociata fra il segnale osservato

$$x(t) = A\sin(\omega t+\phi) + n(t)$$

e il riferimento noto $r(t) = \sin \omega t$

calcolando $R_{xr}(\tau) = E[x(t+\tau)r(t)] = ½ A \cos(\omega\tau +\phi)$

e determinando così sia l'ampiezza A che la fase ϕ incognite.

In pratica la stima della $R_{xr}(\tau)$ verrà fatta su un tempo finito di osservazione T e dunque sarà soggetta a fluttuazioni.
Possiamo calcolarne il contributo, dato che nasce dalla media temporale del prodotto n(t)r(t),
che è evidentemente maggiorata dalla media temporale di n(t)
già considerata in precedenza, per cui la varianza residua è
$\sigma^2 = \sigma_n^2/2BT$

ERRORI DI STIMA DELLE FUNZIONI DI CORRELAZIONE

Vi sono varie definizioni per la stima delle funzioni di correlazione (auto e incrociate) su un tempo di osservazione limitato, fra cui

$$R_{xy}(\tau) = \frac{1}{T-\tau}\int_0^{T-\tau} x(t+\tau)y(t)\,dt \qquad \text{per} \quad 0 \le \tau < T$$

$$R_{xy}(\tau) = \frac{1}{T}\int_0^{T} x(t+\tau)y(t)\,dt \qquad \text{per} \quad 0 \le \tau < T$$

Consideriamo quest'ultima, per cui s'immagina che il tempo di osservazione complessivo sia T+τ
(per il più grande valore di τ che c'interessa)

Come per la precedente, si tratta di una stima non polarizzata, dato che la sua aspettazione coincide con la funzione stessa.

L'errore di stima è rappresentato dalla varianza

$$\sigma^2[R_{xy}(\tau)] = E[(R_{xy}(\tau) - R_{xy}(\tau))^2] = E[R^2_{xy}(\tau)] - R^2_{xy}(\tau) =$$

$$= \frac{1}{T^2}\int_0^T dv \int_0^T du[x(v+\tau)y(v)\,x(u+\tau)y(u) - R^2{}_{xy}(\tau)]$$

dove entrano in gioco i momenti statistici del quarto ordine.

Grane indicibili si risolvono se i processi x(t) e y(t) sono congiuntamente normali (per tali processi tutte le statistiche sono esprimibili in termini di quelle del secondo ordine), nel qual caso l'espressione si semplifica alquanto e si arriva a scrivere:

$$\sigma^2[\hat{R}_{xy}(\tau)] = \frac{1}{T}\int_{-T}^{T}\left(1-\frac{|\xi|}{T}\right)\Big(R_{xx}(\xi)R_{yy}(\xi) + R_{xy}(\xi-\tau)R_{yy}(\xi+\tau)\Big)d\xi$$

Tale espressione che gode dell'importante proprietà di annullarsi per $T \to \infty$ (e dunque la stima è consistente) e di semplificarsi ulteriormente, riducendosi al secondo fattore dell'integrando, per T sufficientemente grandi (rispetto ai ritardi τ che interessano), in tal caso portando all'infinito gli estremi d'integrazione:

$$\sigma^2[\hat{R}_{xy}(\tau)] \approx \frac{1}{T} \int_{-\infty}^{\infty} \left(R_{xx}(\xi) R_{yy}(\xi) + R_{xy}(\xi - \tau) R_{yy}(\xi + \tau)\right) d\xi$$

Nel caso della stima dell'autocorrelazione a ritardo zero l'integrando si riduce a $2R_{xx}(\xi)$. Mentre si riduce a $R_{xx}(\xi)$ per la stima dell'autocorrelazione a grandi valori di τ, nel qual caso si sfrutta il fatto che $R^2_{xx}(\xi) >> R_{xx}(\xi+\tau) R_{xx}(\xi-\tau)$.

Un caso particolare molto importante è quello in cui entrambi i processi x(t) e y(t) sono rumore a media nulla e banda limitata B.
Dalle espressioni precedenti si ottiene allora:

$$\sigma^2[\hat{R}_{xy}(\tau)] \approx \frac{1}{2BT} \left[R_{xx}(0) R_{yy}(0) + R^2{}_{xy}(\tau)\right]$$

Questa trova utile applicazione quando i due processi siano così esprimibili:

$x(t) = s(t) + n(t)$ con $R_{xx}(0) = S + N$
$y(t) = s(t) + m(t)$ con $R_{yy}(0) = S + M$

cioè con rumori non correlati fra loro (e con il segnale), aventi rispettivamente varianza N ed M.

Si tratta di un caso importantissimo, riguardante l'osservazione di uno stesso fenomeno attraverso due diversi sensori, su cui torneremo.

Nel caso specifico dell'estrazione della sinusoide di frequenza nota considerato prima, si ha $y(t) = r(t)$
con $m(t)=0$ dato che il segnale di riferimento è "pulito", e quindi

$R_{xx}(\tau) = R_{ss}(\tau) + R_{nn}(\tau) = \frac{1}{2} A^2 \cos(\omega\tau) + R_{nn}(\tau)$

$R_{yy}(\tau) = \frac{1}{2} \cos(\omega\tau)$

$R_{xy}(\tau) = \frac{1}{2} A \cos(\omega\tau + \phi)$ (fase ϕ intesa rispetto al riferimento)

E la varianza della stima della funzione di correlazione incrociata si calcola con l'espressione precedente di $\sigma^2[\hat{R}_{xy}(\tau)]$

Diversa è la situazione quando il riferimento è rumoroso, come di necessità nel caso seguente.

ESTRAZIONE DI UNA SINUSOIDE

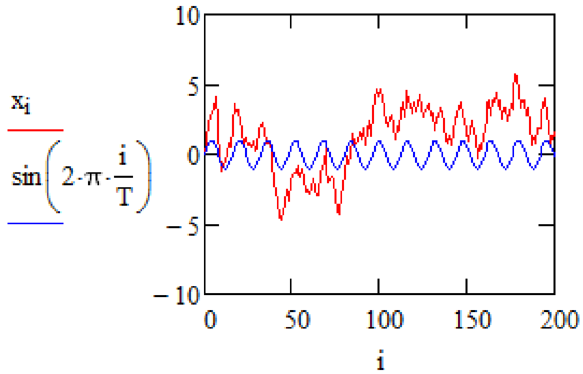
Questa volta non conosciamo la frequenza angolare ω nè l'ampiezza di una sinusoide immersa in rumore. Vogliamo stimare questi parametri.

La soluzione consiste nel calcolare l'autocorrelazione del segnale osservato $x(t) = A\sin(\omega t+\phi) + n(t)$

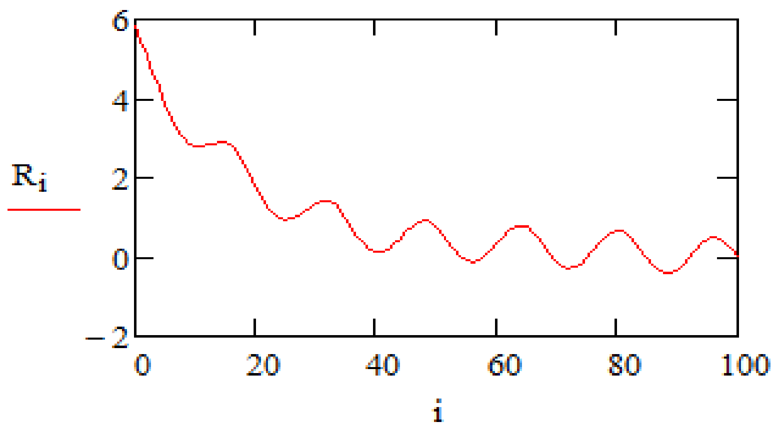
ottenendo $R_{xx}(\tau) = E[x(t+\tau)x(t)] = ½\, A \cos(\omega\tau) + R_{nn}(\tau)$

e consentendo così di determinare le grandezze incognite della sinusoide (ampiezza e frequenza, ma non la fase).

Il grafico a fianco (in rosso) rappresenta una sinusoide di ampiezza unitaria (con periodo T=16 campioni) in presenza di forte rumore, con SNR = 0.1 in energia.



Il grafico sotto rappresenta l'autocorrelazione della sinusoide rumorosa, calcolata con FFT su 4096 campioni.



Quesito.
L'autocorrelazione per ritardo zero vale circa 6. Come si interpreta questo numero?

Anche qui, su un tempo di osservazione finito, vanno consideratgli effetti delle fluttuazioni di stima, certamente maggiori che nel caso considerato prima (sinusoide di frequenza nota).

ANALISI SPETTRALE

Anche l'analisi spettrale può esser fatta rientrare a tutti gli effetti fra le tecniche di correlazione:
infatti il calcolo dei coefficienti di Fourier non è altro che il calcolo della stima delle funzioni di correlazione incrociata (a ritardo zero) fra il segnale analizzato e una opportuna serie di segnali armonici.

Non approfondiamo l'argomento, limitandoci a menzionare alcuni punti essenziali.

1) non vi è generalmente alcun segnale di periodo noto a cui fare riferimento per stabilire la frequenza fondamentale e procedere nel modo indicato sui libri di testo di analisi matematica:

2) si sceglie pertanto un periodo di osservazione T di durata conveniente (assunto come periodo fondamentale) e si esegue il calcolo sui campioni dell'osservazione, che immaginiamo spaziati di δt e dunque in numero totale $n = T/\delta t$

3) si utilizzerà una qualche versione di trasformata discreta basata sulla FFT (fast Fourier transform) (Cooley e Tuckey, 1965) che presenta lo straordinario vantaggio che il numero di operazioni (moltiplicazioni, che sono le più onerose non è proporzionale a n^2 ma a $n \log(n)$; e allora si richiede generalmente che n sia una potenza intera di 2.

4) la prima riga spettrale, come la differenza fra due righe successive cioè la *risoluzione spettrale*, è $\delta f=1/T$, cioè dipende solo dalla durata del tempo di osservazione; mentre la frequenza massima $f_{max} = n/2T$ dipende anche dal numero n dei campioni.

5) elaborando dati relativi a processi stocastici, lo spettro di potenza si ottiene direttamente dal modulo quadro della FFT; in alternativa si calcola prima l'autocorrelazione e poi si trasforma.

6) le stime spettrali di processi casuali ottenute con un singolo spettro sono soggette a forti fluttuazioni, con deviazione standard pari alla stima stessa *indipendentemente dalla lunghezza T del record* (aumentando la quale aumenta il numero delle stime, ma non la loro precisione).

7) per ridurre le fluttuazioni di stima si possono fare delle medie su campioni spettrali adiacenti (smoothing spettrale), ma allora si perde risoluzione; più spesso si suddivide il periodo di osservazione in più parti, si calcolano gli spettri corrispondenti e poi se fa la media degli spettri elementari detti in gergo *periodogrammi*
(anche qui in effetti si perde risoluzione, ma la FFT costa meno).

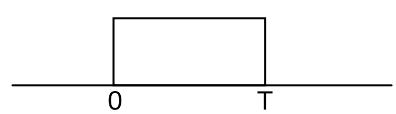
8) nel procedimento anzidetto occorre attenzione qualora abbia interesse la riga spettrale corrispondente a una sinusoide, perchè?

9) un aspetto importante (delicato, ma spesso trascurato) dell'analisi spettrale riguarda la *finestratura* (windowing) dei dati.

Determinare lo spettro di un processo stocastico richiede, in linea di principio, di utilizzare tutta l'informazione disponibile su tutto l'asse dei tempi. In pratica, come già detto, il calcolo si esegue su un tempo di osservazione T, integrando (o sommando) il prodotto fra il processo e il riferimento sinusoidale su tale intervallo.

Questo significa però aver finestrato i dati, cioè averli moltiplicati per una funzione finestra rettangolare, della seguente forma:

$w(t) = u(t) - u(t-T)$

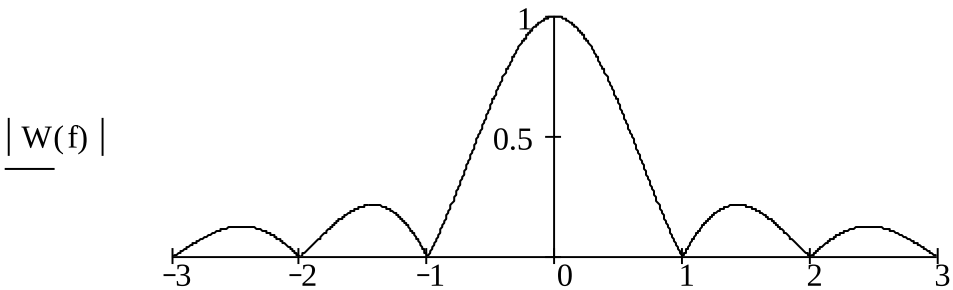


Ciò ha una naturale conseguenza: lo spettro che si stima non rappresenta il processo, ma il prodotto del processo per la finestra w(t), cioè la convoluzione fra lo spettro anzidetto e lo spettro della finestra.

L’andamento della trasformata di Fourier W(f) della finestra è tuttavia ricco di lobi

$$|W(f)| = (\sin \pi f T)/\pi f$$

come mostrato nella figura qui sotto (calcolata per T = 1 s)



I primi lobi laterali sono 13 dB sotto il massimo, mentre il loro inviluppo complessivo decresce con la frequenza secondo la legge
–20 dB/decade (cioè con legge 1/f).

Questo significa, per esempio, che per ogni riga “vera” dello spettro se ne potranno osservare varie altre, spurie (si badi al significato di T), fenomeno che va sotto il nome di leakage spettrale.

Per alleviare il problema è stato proposto l’impiego di finestre di altra forma, in genere con andamento più smussato rispetto a quella rettangolare, alle quali corrisponda quindi un andamento meno oscillante nel dominio della frequenza.

Fra le tante ricordiamo la finestra di Hanning (coseno sollevato o coseno quadrato):

$$w(t) = ½(1- \cos(2\pi t/T)) = 1 – \cos^2(\pi t/T) \qquad \text{fra } 0 \text{ e } T$$

che ha il primo lobo 32 dB sotto al massimo, e inviluppo che decresce con la frequenza secondo la legge –40 dB/decade ($1/f^2$).

Ma si paga un prezzo: il lobo principale si allarga, ciò che significa una riduzione della risoluzione spettrale.

Un punto delicato riguarda l'impiego delle tecniche di analisi spettrale per la ricerca di deboli segnali periodici (di ampiezza e frequenza incognite) in presenza di rumore.

Qui la durata dell'intervallo di osservazione su cui calcolare gli spettri non è indifferente, come lo è invece per la stima del contributo spettrale incoerente di un processo stocastico.

Sfruttare la coerenza di un segnale armonico, per migliorare il rapporto segnale rumore o semplicemente per consentirne l'individuazione, richiede infatti di eseguire FFT sull'intervallo più lungo possibile. Durante il quale i contributi della riga si sommeranno coerentemente.

Mentre operando su intervalli più brevi e poi sommando gli spettri si perde il vantaggio della coerenza di fase del segnale, a meno di non prendere misure opportune a tal fine.

Trasformate su periodi lunghi, d'altra parte, possono risultare non attuabili
- perchè il numero di campioni in gioco, nonostante l'algoritmo veloce, eccede le possibilità di calcolo
- per motivi di altra natura (come l'effetto Doppler dovuto ai moti della Terra rispetto alle sorgenti, nella ricerca di onde gravitazionali periodiche emessa da pulsar o da oggetti consimili).

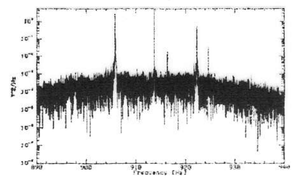
Allora la soluzione consiste in....

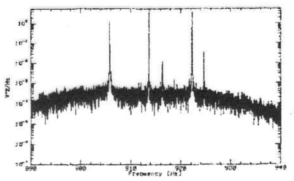
Spettri sperimentali del segnale d'uscita di un rivelatore gravitazionale risonante (Explorer al CERN).

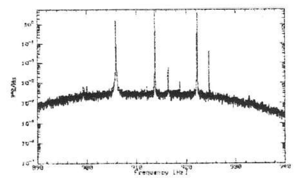
Si notano due picchi corrispondenti ai due modi meccanici del rivelatore (sbarra sensibile e trasduttore risonante), la riga corrispondente al segnale di calibrazione usato per monitorare il guadagno dell'amplificatore SQUID
e altre righe spurie: rete a 900 Hz, risonanze impreviste, ecc.

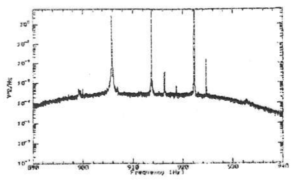
Gli spettri sono stati presi in condizioni di funzionamento stazionario:
nell'ordine un solo spettro,
medie su 5 periodogrammi,
medie su 25, medie su 125

Si nota la fortissima variabilita' delle stime a bassa statistica, tenendo presente che il fondo è approssimativamente costante nel range osservato









In tutti i grafici, la scala orizzontale, che si legge assai male, si estende fra 890 Hz e 940 Hz

correlazione 10 bis

IL FILTRO ADATTATO

Consideriamo ora il **filtro adattato** (*matched filter*),
che è usato per la rivelazione e la stima dei parametri di segnali transitori di forma nota, ed è basato su una operazione di convoluzione del tutto equivalente a una di correlazione.

Questa volta l'osservazione contiene (o non contiene) un segnale di forma nota s(t) , ma di ampiezza A e tempo di occorrenza t_o entrambi sconosciuti:

$$x(t) = A\,s(t-t_o) + n(t)$$

L'obiettivo NON è estrarre il segnale dal rumore stimandone l'andamento nel tempo, cioè ricostruire il segnale, ma individuarne la presenza ed eventualmente stimarne i parametri incogniti.

Per questo correliamo l'osservazione con la forma d'onda nota s(t), integrando dunque in dt il prodotto $x(t+\tau)s(t)$ oppure, che è lo stesso, $x(t)s(t-\tau)$. Questa operazione darà luogo a un picco di risposta, proporzionale ad A (trascurando il contributo del rumore), per $\tau = -t_o$.

Anzichè correlare l'osservazione x(t) con il riferimento s(t), possiamo convolverla con una funzione h(t), integrando allora in $d\alpha$ il prodotto $x(\alpha)h(t-\alpha)$, cioè usando un filtro lineare con risposta impulsiva h(t):

$$m(t) = \int_{-\infty}^{t} h(t-\alpha)\,x(\alpha)\,d\alpha = \int_{0}^{\infty} h(\alpha)\,x(t-\alpha)\,d\alpha$$

Dove evidentemente si ottiene lo stesso risultato del correlatore se

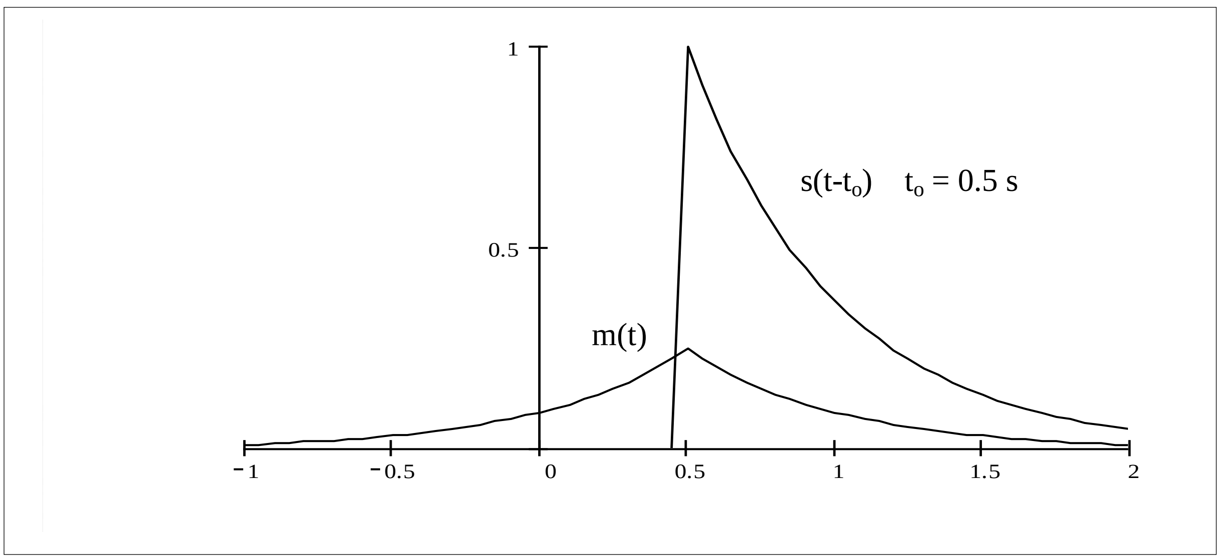
$$h(t) = s(-t)$$

cioè la risposta impulsiva coincide con la forma d'onda nota invertita secondo l'asse dei tempi.

Il filtro così ottenuto è chiamato *filtro adattato*, in quanto adattato alla forma del segnale. La sua funzione di trasferimento $H(j\omega)$ è la trasformata di Fourier di $h(t) = s(-t)$, cioè $H(j\omega) = S^*(j\omega)$

ESEMPIO

Uscita del filtro adattato a un segnale esponenziale in risposta al segnale applicato al tempo $t_o = 0.5$ s. Si nota che la risposta non riproduce affatto il segnale, e presenta un picco proprio a 0.5 s.



Nota storica

La teoria del filtro adattato (come del resto quella del filtro ottimo di Wiener) è stata sviluppata sostanzialmente durante l'ultima guerra mondiale, per la rivelazione di segnali radar

(segnali evidentemente di forma nota, di cui interessa sopratutto il ritardo in ricezione, per stabilire così la distanza del bersaglio).

Si trattava dunque di materiale classificato, che è stato pubblicato solo più tardi. Per esempio nel lavoro di B.M. Dwork *Detection of a pulse superimposed on fluctuation noise*, Proc. IRE, luglio 1950.

Si dimostra che il filtro con le caratteristiche dette (se il rumore è bianco) fornisce il massimo valore del rapporto segnale/rumore, cioè è ottimo, fra tutti i possibili filtri lineari. Si dimostra anche che se il rumore ha distribuzione normale il filtro adattato lineare è anche ottimo rispetto a qualsiasi filtro nonlineare.

Specifichiamo meglio cosa s'intenda, in questo caso, per rapporto segnale/rumore dopo il filtraggio.

Per *segnale* qui intendiamo il quadrato del valore massimo del segnale m(t) all'uscita del filtro, cioè

$m^2(t_o)$

Per *rumore* intendiamo la varianza σ^2 del rumore all'uscita del filtro

Dunque abbiamo $SNR = m^2(t_o)/\ \sigma^2$

Se il filtro ha funzione di trasferimento $H(j\omega)$, il segnale d'uscita, per l'ingresso $As(t\text{-}t_o)$, sarà l'antitrasformata del prodotto $A\,H(j\omega)\,S(j\omega)\exp(-j\omega t_o)$, mentre la varianza del rumore sarà data dall'integrale dello spettro del rumore d'uscita, cioè:
$S_n(\omega)\,|H(j\omega)|^2$

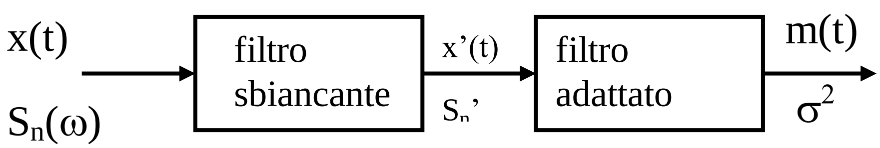
Il rapporto segnale rumore sarà dunque, per un segnale di ampiezza A applicato al tempo generico t_o:

$$SNR = \frac{A^2}{2\pi}\frac{\left|\int_{-\infty}^{\infty} e^{-j\omega t_o} S(j\omega) H(j\omega)\, d\omega\right|^2}{\int_{-\infty}^{\infty} \left|H(j\omega)\right|^2 S_n(\omega)\, d\omega}$$

Se il rumore è bianco, cioè $S_n(\omega)$ = cost, si trova che SNR è massimo per $H(j\omega) = S^*(j\omega)$, assumendo il valore

$$SNR = \frac{A^2}{2\pi S_n} \int_{-\infty}^{\infty} \left| S\left(j\omega \right) \right|^2 d\omega = \frac{A^2}{S_n} \int_{-\infty}^{\infty} \left| s\left(t \right) \right|^2 dt$$

Se il rumore non è bianco, allora va reso tale con un *filtro sbiancante*



Come si procede?
Grazioso esercizio che viene affidato alla sagacia del Lettore di queste note. Si badi comunque che lo sbiancamento del rumore influenzerà anche il segnale....

Avendolo risolto, si troverà che il rapporto segnale/rumore è dato dall'espressione (avendo posto A=1):

$$SNR = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\left| S(j\omega) \right|^2}{S_n(\omega)} d\omega$$

REALIZZAZIONI DI CORRELATORI

Il calcolo di una funzione di auto (o cross) correlazione viene oggi generalmente eseguito al calcolatore su sequenze di dati discretizzati nel tempo, di durata opportuna ai fini della riduzione delle fluttuazioni di stima.

Spesso risulta meno costoso, in termini di tempo di calcolo, eseguire prima la FFT dei dati, poi calcolare il modulo quadro dello spettro (o delle medie di più periodogrammi) e poi calcolare la FFT inversa per ottenere la funzione di correlazione.

In passato si utilizzavano invece correlatori hardware, una soluzione che può risultare ancora oggi valida in situazioni particolari, per esempio quando si tratti di segnali di altissima velocità, con larghezza di banda molto grande.

Che creano problemi, sia di conversione A/D che di elaborazione, sopratutto se c'e' l'esigenza di ottenere risultati in tempo reale.

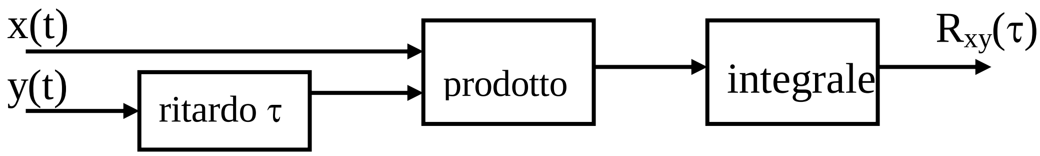
Allora si tratta di realizzare in forma analogica, oppure digitale con hardware dedicato, l'operazione espressa dalla formula

$$R_{xy}(\tau) = \frac{1}{T}\int_0^T x(t+\tau)\,y(t)\,dt$$

dove le operazioni essenziali sono quelle di *ritardo*, *prodotto* fra due segnali e *integrazione* nel tempo.

Le soluzioni sono evidentemente diversissime (e le più varie)
a seconda che la realizzazione sia analogica oppure digitale

Schema di principio per la realizzazione di un correlatore



con tanti elementi di ritardo, di moltiplicazione e d'integrazione quanti sono i punti della funzione che si vogliono calcolare.

Un elemento critico (in termini di realizzazione nel caso analogico, di velocità di calcolo in quello digitale) è il moltiplicatore.

Per questo è stata proposto, e usato in varie applicazioni, l'impiego di segnali a pochi livelli di quantizzazione, fino al caso estremo, nei *correlatori a polarità*, di quantizzazione con un solo bit.
Dove il moltiplicatore diventa un oggetto semplicissimo: un circuito logico che dia in uscita "1 logico" quando gli ingressi sono dello stesso segno, altrimenti "0 logico" (si può ridurlo a una porta logica AND).

In questo caso si perde evidentemente qualsiasi informazione sull'intensità dei segnali, ma il correlatore fornisce comunque un risultato utile. Si dimostra infatti che l'uscita del correlatore a polarità fornisce la seguente informazione:

$R(\tau) = (2/\pi) \arcsin(c_{xy}(\tau))$

dove $c_{xy}(\tau)$ è la funzione di covarianza normalizzata. Per valori piccoli dell'argomento rispetto all'unità, la formula precedente fornisce proprio la covarianza normalizzata fra i due processi analizzati.

Lo schema di dimostrazione è il seguente. Si considerano due processi normali, esprimendone la densità congiunta (normale anch'essa) in funzione della covarianza normalizzata. E si calcola la funzione di correlazione incrociata usando la definizione di media d'insieme, in cui interviene appunto la densità congiunta, ma considerando segnali sottoposti a quantizzazione a due livelli. Integrando ....

Nei correlatori quantizzati si paga un prezzo anche dal punto di vista del rapporto segnale rumore (fluttuazione residua a parità di tempo d'integrazione), che peggiora riducendosi al 64% rispetto al correlatore ideale di riferimento (88% nel caso di segnali rappresentati con due bit).

Questa tecnica, introdotta nel 1944 (J.H. Van Vleck, D. Middleton *The Spectrum of Clipped Noise*, Proc. IEEE, gennaio 1966) e descritta da S. Weinreb (tesi di dottorato al MIT del 1963), ha trovato impiego, in radioastronomia, fisica medica, e altri settori.

Uno schema appena più complicato prevede l'aggiunta di dither (rumore a densità uniforme, ma anche un'onda triangolare, anch'essa a densità uniforme) ai due processi: in tal caso la precedente formula viene linearizzata, cioè si ottiene direttamente la covarianza normalizzata (A.R. Meo *Error probabilities and estimate variances of E1B correlators* Alta Frequenza, maggio 1969).

TECNICA DI TWISS-BROWN

Avendo a disposizione due sensori di un medesimo segnale, con rumori indipendenti fra loro, un’ovvia possibilità e’ quella di sommare le due uscite, migliorando allora il rapporto segnale rumore di un piccolo fattore (2 in energia). E ottenendo una stima dell’andamento del segnale in funzione del tempo.

Assai più potente in termini di miglioramento del rapporto segnale rumore, al prezzo però di perdere informazione sull’andamento del segnale, stimandone soltanto il valore quadratico medio,
è la tecnica a correlazione incrociata proposta da R.H.Brown e R.Q.Twiss (*Correlation between photons in two coherent beams of light* Nature, gennaio 1956, pag.27).

Successivamente impiegata (a volte riscoperta) da vari altri autori, in radioastronomia, in termometria di rumore, per misure del campo magnetico interplanetario, ecc.

Rappresentando i segnali dei due sensori nella forma di processi stocastici stazionari

$x(t) = s(t) + n(t)$ con $R_{xx}(0) = S + N$ $R_{sn}(\tau) = 0$
$y(t) = s(t) + m(t)$ con $R_{yy}(0) = S + M$ $R_{sm}(\tau) = 0$ $R_{nm}(\tau) = 0$

risulta ovvio che la correlazione incrociata fornisce la varianza del segnale: $R_{xy}(0) = S$

Va poi considerato l’errore di stima, ma un punto essenziale è che veramente non vi sia alcuna correlazione fra i rumori associati alle osservazioni dei due sensori.

La varianza residua di stima si ottiene dall’espressione trovata prima per rumore a banda limitata B

$$\sigma^2[\hat{R}_{xy}(\tau)] \approx \frac{1}{2BT}\left[R_{xx}(0)R_{yy}(0) + R^2{}_{xy}(\tau)\right]$$

Nel caso che ci interessa abbiamo allora

$$\sigma^2[\hat{R}_{SS}(0)] \approx \frac{1}{2BT}\left[(S+N)(S+M)+S^2\right]$$

che in condizioni di basso rapporto segnale/rumore e per N=M diventa

$$\sigma^2[\hat{R}_{SS}(0)] \approx \frac{N^2}{2BT}$$

La cui radice quadrata rappresenta la deviazione standard dell'errore di fluttuazione nella stima della varianza del segnale.

Ecco come per tempi d'integrazione sufficientemente lunghi si possa, per esempio, fare della termometria di rumore di altissima precisione (errori inferiori a 0.5 mK) alla temperatura dell'elio liquido pur usando un preamplificatore a FET con temperatura di rumore attorno a 100 mK (H.H.Klein, G.Klempt, L.Storm *Measurement of the Thermodynamic Temperature of* ^{4}*He at Various Vapour Pressures by a Noise Thermometer*, Metrologia, 1979, vol.15, pag. 143).

Si noti che il fattore di riduzione del rumore dipende fortemente (qui come del resto in altri casi) dalla larghezza di banda del rumore.
Le misure più difficili (nel senso che richiedono tempi più lunghi) sono dunque quelle in presenza di rumore a banda stretta: sia in continua che nel caso di sistemi risonanti ad alto Q. Ma vedremo in seguito una tecnica per aggirare anche questo problema.

COINCIDENZE

Fre le tecniche di correlazione rientra il metodo delle **coincidenze,** largamente usato in molti settori della fisica, che fu introdotto da Bruno Rossi per lo studio dei raggi cosmici (circuito di coincidenza alla Rossi, impiegante due triodi per realizzare quella che oggi chiamiamo porta AND).

Si tratta, più precisamente, di *correlazione impulsiva* fra eventi. Perchè ora non consideriamo segnali continui nel tempo, ma *eventi impulsivi*, non caratterizzati da una forma d'onda, ma soltanto da un tempo di occorrenza e da una intensità (ampiezza massima, energia totale, ecc.).

Il metodo delle coincidenze viene generalmente usato per individuare effetti comuni che si manifestano in due (o più) diversi rivelatori di uno stesso fenomeno, quando è presente un rumore di fondo tale da impedire l'osservazione del fenomeno usando i segnali di un solo rivelatore.

Il modello più semplice è il seguente.
Si hanno due rivelatori, A e B, nei quali per effetto del rumore si generano eventi con frequenza media λ_A e λ_B , mentre la causa comune si manifesta con eventi aventi frequenza media λ_C
molto inferiore a λ_A e $\lambda_{B.}$ Tutti con distribuzione di Poisson nel tempo.

In un periodo di osservazione T
il rivelatore A registra mediamente $\lambda_A T$ eventi dovuti al fondo ed
$\lambda_C T$ eventi dovuti alla causa comune,
mentre il rivelatore B vede $\lambda_A T$ eventi dovuti al fondo e gli stessi $\lambda_C T$ eventi comuni.

E' evidente che il numero di eventi registrati dai rivelatori (n_A e n_B) è soggetto a fluttuazione, cioè n_A è uguale a $\lambda_A T + \lambda_C T$ solo in media, e così pure n_B è uguale a $\lambda_B T + \lambda_C T$ solo in media.

Usando un solo rivelatore, non si può stabilire nulla a meno che il segnale sia così frequente che, considerando per esempio il rivelatore A, lo scarto fra l'osservazione n_A e il valore aspettato $\lambda_A T$ sia statisticamente significativo,
diciamo almeno dell'ordine della radice quadrata dell'osservazione stessa (rivelazione a un sigma).

La sensibilità aumenta grandemente se si considerano i due rivelatori come parte di un esperimento congiunto, considerando significativi solo gli eventi registrati da essi contemporaneamente, cioè in coincidenza temporale.

Questa operazione di coincidenza, che è l'equivalente del calcolo della funzione di correlazione incrociata fra le uscite dei due rivelatori, si attua scegliendo una *finestra temporale di coincidenza* T_w, e contando contando il numero delle *coincidenze*, cioè il numero di volte che nei due rivelatori si registrano eventi che distano fra loro meno di T_w.

La scelta della finestra è tutt'altro che banale, dato che da essa dipende la sensibilità finale dell'esperimento di coincidenza. Ad essa concorrono vari elementi: di natura fisica e di natura tecnica.

Il fenomeno comune si manifesta veramente in perfetto sincronismo nei due rivelatori o ci può essere qualche effetto di ritardo (come il tempo luce, fra due rivelatori gravitazionali a distanza)?

Qual è la durata intrinseca dei segnali impulsivi dei rivelatori? Quale la accuratezza degli orologi dei due rivelatori? E via dicendo.

Altri problemi, a monte di quanto detto, riguardano la definizione stessa di *evento*, inteso generalmente come un segnale che supera una soglia (predeterminata dallo sperimentatore o aggiustata poi al meglio?).
Che generalmente non crea problemi quando si tratta di segnali grandi rispetto al fondo (l'uscita di un rivelatore di particelle). Mentre crea problemi quando i segnali sono appena distinguibili dal fondo.

Valutiamo ora la sensibilità di un esperimento di coincidenza, calcolando innanzitutto il numero aspettato delle *coincidenze accidentali*, dovute ai fondi dei due rivelatori che assumiamo indipendenti fra loro.

In un generico intervallo Δt di piccola durata, la probabilità di trovare un evento di fondo del rivelatore A è $p_A=\lambda_A \Delta t$, quella di trovarne uno del rivelatore B, $p_B=\lambda_B \Delta t$. Data l'indipendenza, la probabilità di trovarne uno di A e uno di B, è $p_A\ p_B$.

La frequenza media delle coincidenze casuali è dunque

$$\lambda = p_A\ p_B/\Delta t = \lambda_A\ \lambda_B\ \Delta t$$

e il loro numero aspettato nell'intervallo T è

$$E[n] = \lambda_A\ \lambda_B\ \Delta t\ T$$

che possiamo anche esprimere, approssimativamente, in termini degli eventi registrati dai due rivelatori
(assumendo cioè $\lambda_A \approx n_A/T$, $\lambda_B \approx n_B/T$)

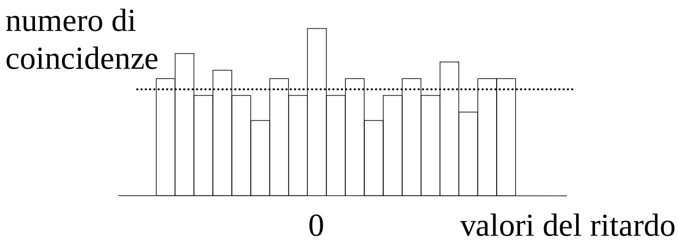
$$E[n] \approx n_A n_B\ \Delta t\ /T$$

notiamo peraltro che l'approssimazione è molto buona se i numeri n_A e n_B sono molto maggiori dell'unità.

Nel risultato dell'analisi in coincidenza, al contributo delle casuali si somma quello delle coincidenze "vere" $E[n_C] = \lambda_C\ T$
che può essere stimato solo per sottrazione del precedente.

Il contributo delle casuali, peraltro, non è noto. Si stima in due modi:
per via teorica, utilizzando la precedente formula $E[n] \approx n_A n_B\ \Delta t\ /T$
(che peraltro è valida solo per processi stazionari)
per via sperimentale, ripetendo più volte l'esperimento di coincidenza, ma ritardando opportunamente gli eventi di un rivelatore rispetto all'altro, e mediando il risultato ottenuto.

Spesso, in pratica, i risultati di un'analisi in coicidenza si presentano nella forma di un *time delay histogram*, cioè graficando i numeri delle coincidenze trovate in corrispondenza di un certo numero di valori del ritardo fra i dati dei due esperimenti.



Il valor medio delle coincidenze per ritardi diversi da zero rappresenta un'ottima stima del numero delle coincidenze accidentali. La differenza fra il numero trovato a ritardo zero e il numero delle casuali, una buona stima del numero di coincidenze "vere", cioè dovute a un fenomeno comune che agisce sui due rivelatori.

Il significato statistico del risultato si può esprimere in termini di

- rapporto critico, dividendo l'eccesso a ritardo zero per la radice quadrata del numero delle casuali;
- probabilità che l'eccesso sia dovuto al caso, calcolando allora la probabilità che una variabile Poissoniana con valor medio pari al numero della casuali assuma valore maggiore o uguale a quello effettivamente trovato a ritardo zero.

Ma questa probabilità caratterizza effettivamente il risultato finale solo se ....

ESEMPIO

I risultati dell'analisi in coincidenza di eventi impulsivi registrati dai rivelatori gravitazionali Explorer (Cern) e Nautilus (Frascati) durante 701 ore di funzionamento forniscono i seguenti dati:

coincidenze a ritardo zero:	19
numero medio delle casuali stimato su 10^4 ritardi:	10.8
probabilità di Poisson:	0.015

Bibliografia

L.E. Franks, *Signal theory*, Prentice Hall, 1969

A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw Hill, 1991

J.S. Bendat, A.G. Piersol, *Random Data – Analysis and Measurement Procedures*, Wiley, 1986

A.D. Whalen, *Detection of signals in noise*, Academic Press, 1971

G. E. Turin, *An Introduction to Matched Filters,* IRE Trans. Information Theory, 1960, pag. 311.

E. R. Davies, *Electronics, Noise and Signal Recovery,* Academic Press, 1993.

T.H.Wilmshurst, *Signal Recovery fron Noise in Electronic Instrumentation*, Adam Hilger, 1985.

L.A.Wainstein, V.D.Zubakov, *Extraction of Signals from Noise* Prentice Hall, 1962.

P. Astone et al., *The fast matched filter for gravitational-wave data analysis* Il Nuovo Cimento, vol. 20 C, n.1, Gennaio 1997, pp. 9-60.