Lezione 1

20 Settembre ore 14-16

Introduzione al corso
Esempi di applicazioni del Signal Processing
Richiami sulla Z-trasformata e sulla DTFT
Progetto di filtri mediante piazzamento poli-zeri
Esempi

Lezione 2

22 settembre 2021 ore 9-11

Progetto FIR con il metodo della serie di Fourier
Filtri FIR mediante campionamento in frequenza: filtri simmetrici e anti-simmetrici a fase lineare
Progetto filtri da prototipi analogici
Tecniche di invarianza all'impulso, al gradino e a alla rampa

Lezione 3

lunedì 27 settembre 2021 14.12

$$H_a(s) \longrightarrow H(z)$$

inv. all'impulso

$$H(z) = \mathcal{Z}\left[\mathcal{L}^{-1}[H_a(s)]\Big|_{t=nT}\right]$$

inv. al gradino

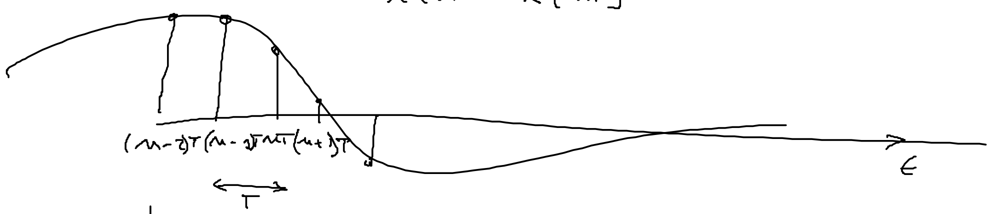
$$H(z) = (1-z^{-1})\,\mathcal{Z}\left[\mathcal{L}^{-1}\left[\frac{H_a(s)}{s}\right]\Big|_{t=nT}\right]$$

inv. alla rampa

$$H(z) = (z\ \ldots\ )\,\mathcal{Z}\left[\mathcal{L}^{-1}\left[\frac{H_a(s)}{s^2}\right]\Big|_{t=nT}\right]$$

$$H_a(s) = \frac{\sum \alpha_k s^k}{\sum \beta_k s^k} \longrightarrow \sum_k \beta_k \frac{d y^k}{d t^k} = \sum_k \alpha_k \frac{d^k x}{d t^k}$$

$x(t)$ $\quad x[nT]$



$$x'(nT) \longrightarrow \frac{x[nT] - x[(n-1)T]}{T}$$

$x(t) \rightarrow \boxed{s} \rightarrow x'(t)$ $\qquad$ $x[n] \rightarrow \boxed{D(z)} \rightarrow x'[n]$

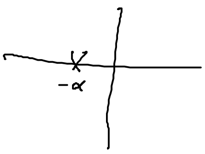
$$\frac{X(z) - z^{-1}X(z)}{T} \qquad D(z) = \frac{1-z^{-1}}{T}$$

$$H(z) = H_a(s)\Big|_{s = \frac{1-z^{-1}}{T}}$$

Esempio

$$H_a(s) = \frac{A}{s+\alpha} \qquad H(z) = \frac{A}{\frac{1-z^{-1}}{T}+\alpha} = \frac{AT}{1-z^{-1}+\alpha T} = \frac{AT}{(1+\alpha T) - z^{-1}}$$

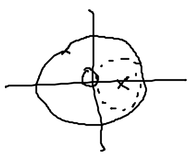
$\alpha > 0$

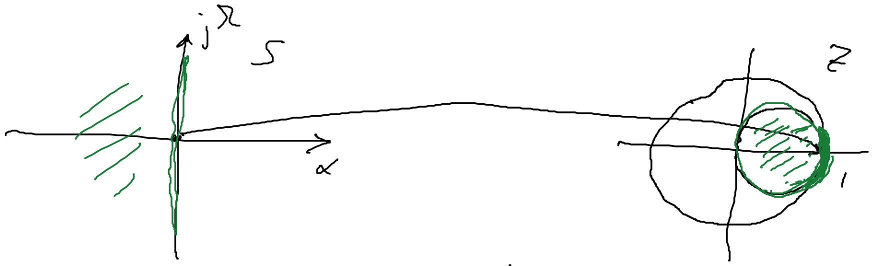


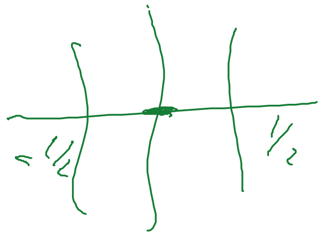
$$y[n] = \frac{1}{(1+\alpha T)}y[n-1] + \frac{AT}{1+\alpha T}x[n]$$

$$= \frac{AT}{1+\alpha T} \cdot \frac{1}{1-\frac{1}{(1+\alpha T)}z^{-1}}$$

$$= \frac{AT}{1+\alpha T} \quad \frac{z}{z-\frac{1}{(1+\alpha T)}}$$



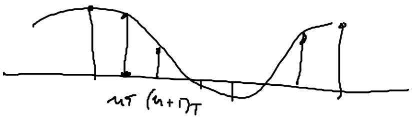




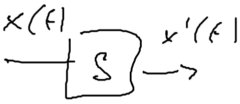
$$S = \frac{1 - z^{-1}}{T}$$

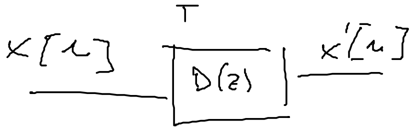
$$ST = 1 - z^{-1} \Rightarrow 1 - ST = z^{-1} \Rightarrow z = \frac{1}{1 - ST}$$

$$z = \frac{1}{1 - \alpha T - j\Omega T} \xrightarrow[\alpha = 0]{} z = \frac{1}{1 - j\Omega T}$$



$$x'(nT) \simeq \frac{x[(n+1)T] - x[n]}{T}$$





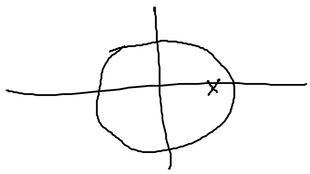
$$D(z) = \frac{z - 1}{T}$$

$$H(z) = H_a(s)\Big|_{s = \frac{z-1}{T}}$$

Esempio

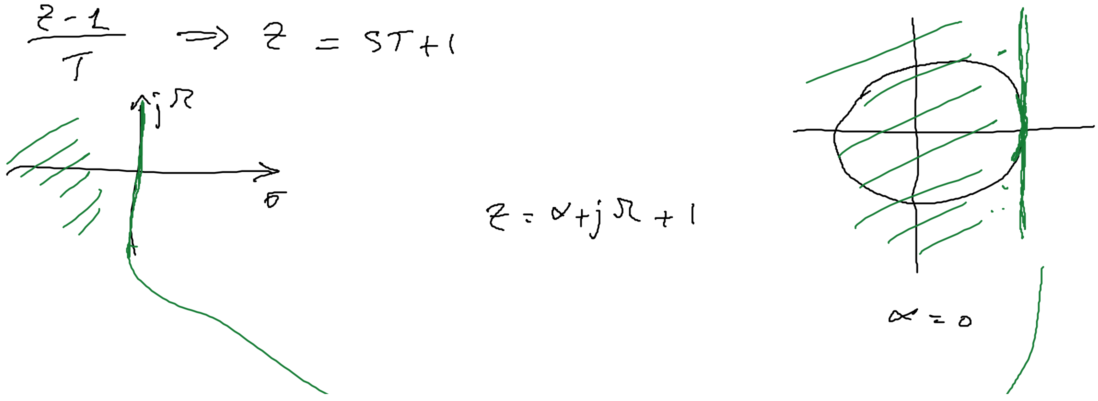
$$H_a(s) = \frac{A}{s + \alpha} \qquad H(z) = \frac{A}{\frac{z-1}{T} + \alpha} = \frac{AT}{z - 1 + \alpha T} = \frac{AT}{z - (1 - \alpha T)}$$

$$= \frac{AT \; z^{-1}}{1 + (\alpha T - 1) z^{-1}}$$

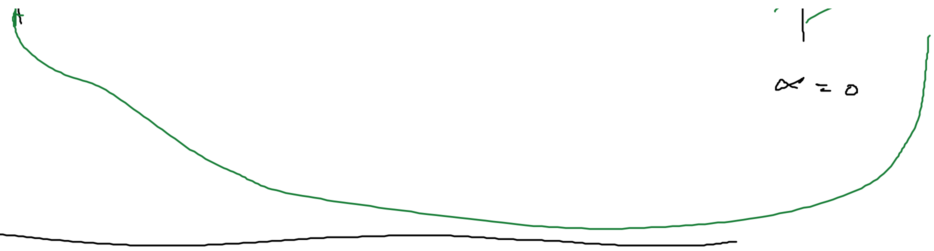


$$y[n] = (\alpha T - 1)\, y[n-1] + AT\, x[n-1]$$

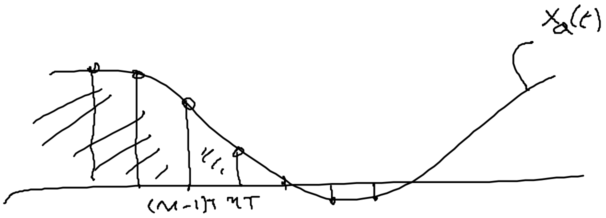
$$S = \frac{z - 1}{T} \Rightarrow z = ST + 1$$



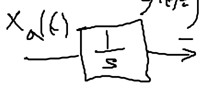
$$z = \alpha + j\Omega + 1$$



$\alpha = 0$

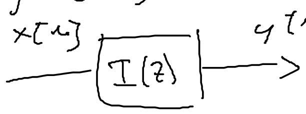


$x[n] = x_a(nT)$



$$y(t) = \int_{-\infty}^{t} x_a(\xi)\, d\xi$$

$$y[n] = y[n-1] + \left(x[n-1] + x[n]\right)\frac{T}{2}$$



$$Y(z) = z^{-1}Y(z) + \frac{T}{2}(z^{-1}+1)X(z)$$

$$\frac{1}{s} \longleftrightarrow I(z) = \frac{Y(z)}{X(z)} = \frac{T}{2}\,\frac{1+z^{-1}}{1-z^{-1}}$$

$$H(z) = H_a(s)\Big|_{s = \frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}}$$

TRASFORMAZIONE BILINEARE

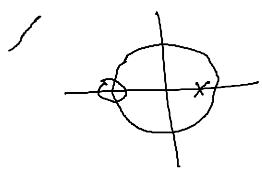
Esempio

$$H_a(s) = \frac{A}{s+\alpha} \rightarrow H(z) = \frac{A}{\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} + \alpha} = \frac{AT(1+z^{-1})}{2(1-z^{-1}) + \alpha T(1+z^{-1})}$$

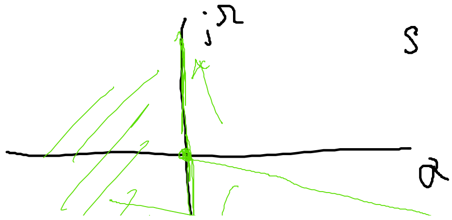
$$= AT\,\frac{1+z^{-1}}{2+\alpha T - (2-\alpha T)z^{-1}} = \frac{AT}{2+\alpha T}\,\frac{1+z^{-1}}{1 - \frac{2-\alpha T}{2+\alpha T}z^{-1}}$$

$$y[n] = \frac{2-\alpha T}{2+\alpha T}\, y[n-1] + AT\left(x[n] + x[n-1]\right)$$

$$= \frac{AT}{2+\alpha T}\,\frac{z+1}{z - \frac{2-\alpha T}{2+\alpha T}}$$



$$s = \frac{2}{T}\,\frac{1-z^{-1}}{1+z^{-1}} \qquad z = -\frac{sT+2}{sT-2} = -\frac{\frac{sT}{2}+1}{\frac{sT}{2}-1}$$

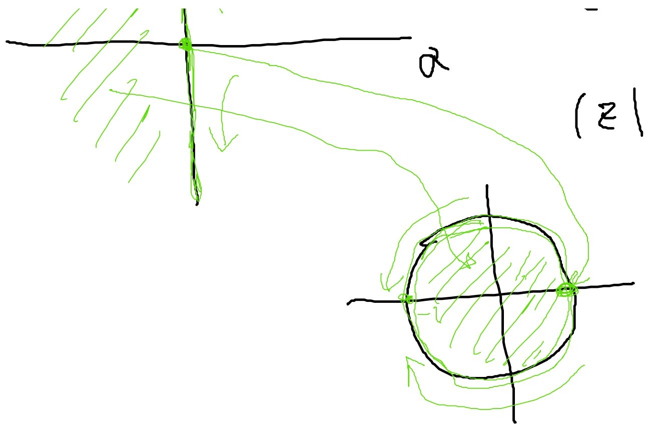


s

$$s = j\Omega$$

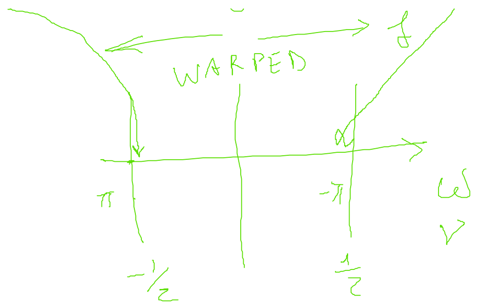
$$z = -\frac{j\Omega T + 2}{j\Omega T - 2}$$

$|z|^2 \qquad \Omega^2T^2+4$



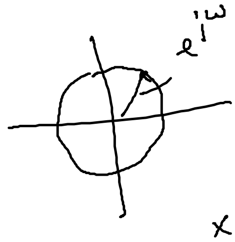
$$j\Omega T - 2$$

$$|z|^2 = \frac{\Omega^2 T^2 + 4}{\Omega^2 T^2 + 4}$$



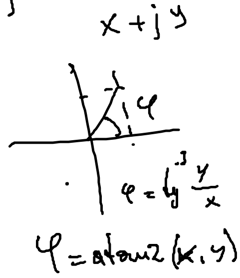
$\Omega = 2\pi f$ rad/sec.

$\omega = 2\pi\nu$ rad



$$z = -\frac{sT+2}{sT-2}$$

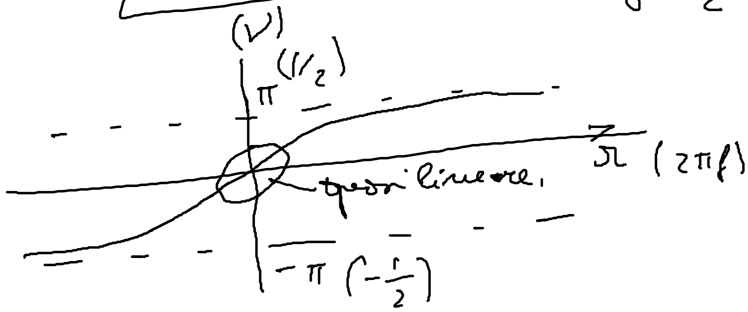
$$e^{j\omega} = -\frac{j\Omega T + 2}{j\Omega T - 2} \qquad e^{j\omega} = -\frac{(j\Omega T+2)(-j\Omega T-2)}{\Omega^2 T^2 + 4}$$



$$e^{j\omega} = \frac{(2+j\Omega T)^2}{\Omega^2 T^2 + 4}$$

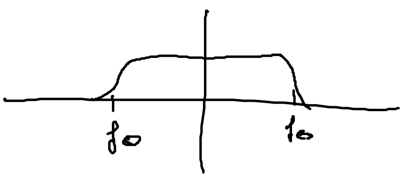
$$\omega = 2\angle\, 2 + j\Omega T = 2\, tg^{-1}\frac{\Omega T}{2}$$

$$\Omega = \frac{2}{T} tg \frac{\omega}{2}$$

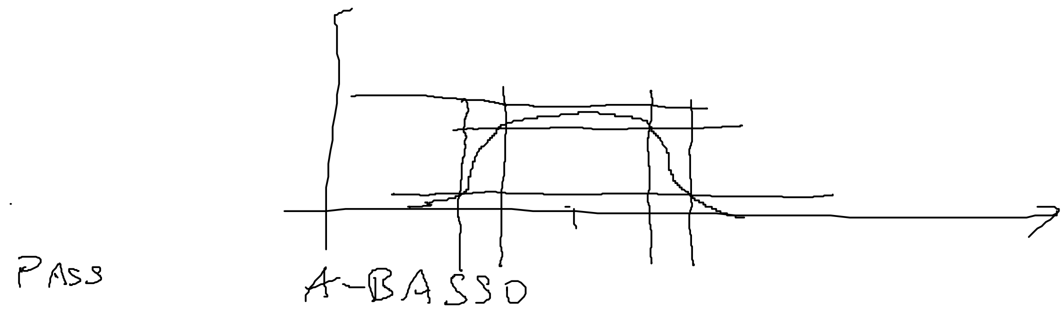


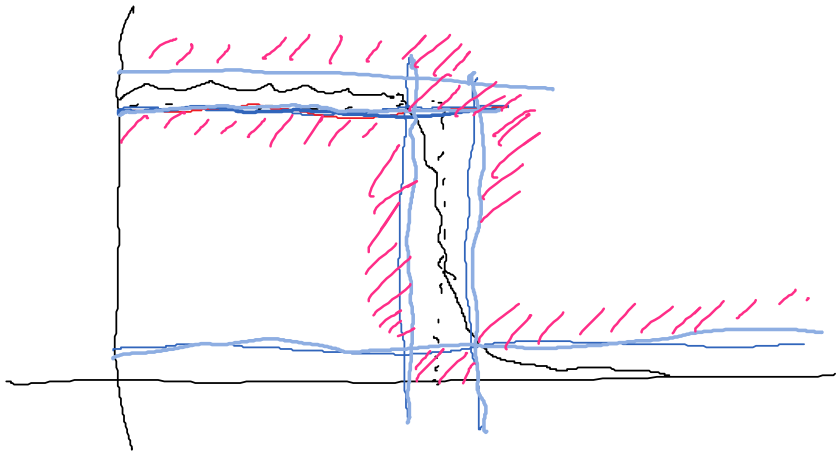
$$\cancel{2}\pi\nu = \cancel{2}\, tg^{-1}\frac{\cancel{2}\pi f T}{\cancel{2}}$$

$$\nu = tg^{-1}\pi f T$$



$$T = \frac{1}{f_c}$$





Filtro di Butterworth (delle dispense)

- Determino N

$$H_{LP}(s)\,H_{LP}(-s) = \frac{1}{1+\left(\frac{-s^2}{\Omega_c^2}\right)^N}$$

$$1+\left(\frac{-s^2}{\Omega_c^2}\right)^N = 0 \Rightarrow \left(-\frac{s^2}{\Omega_c^2}\right)^N = -1$$

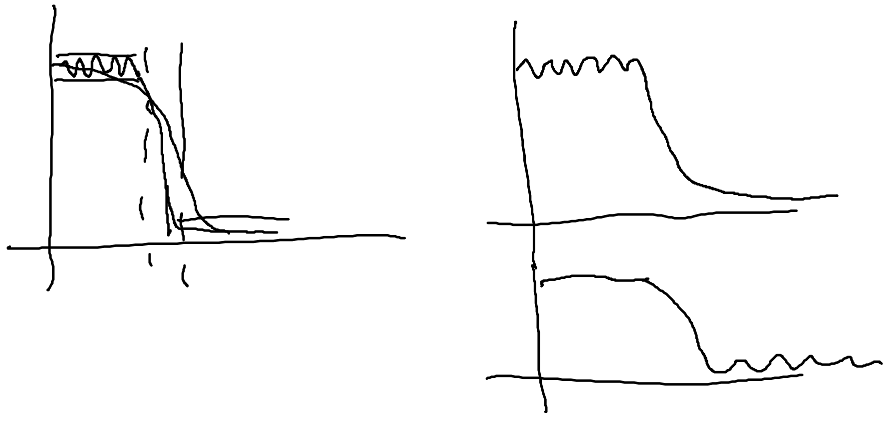
$$(-1)^N s^{2N} = \Omega_c^{2N}(-1)$$

$$s^{2N} = \Omega_c^{2N}(-1)^{1+N}$$

$$\rho^{2N} e^{j\theta 2N} = \Omega_c^{2N} e^{j\pi(1+N)+k2\pi} \Rightarrow \qquad \rho = \Omega_c$$

$$\theta 2N = \pi(1+N) + k2\pi$$

$$\theta = \frac{\pi(1+N)}{2N} + \frac{k\pi}{N}$$



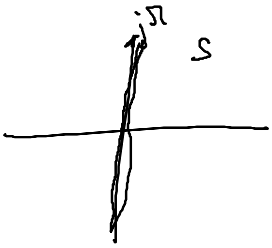
TRASF. PASSA-BASSO (Analogico) $\longrightarrow$ Filtro discreto (analogico)

Dagli appunti.

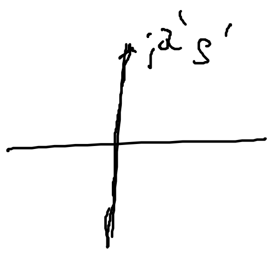
Esempio

PASSA BASSO Ω_c PASSA ALTO Ω_c'

Ω Ω'

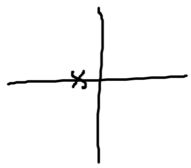


$$s = \frac{\Omega_c \Omega_c'}{s'}$$

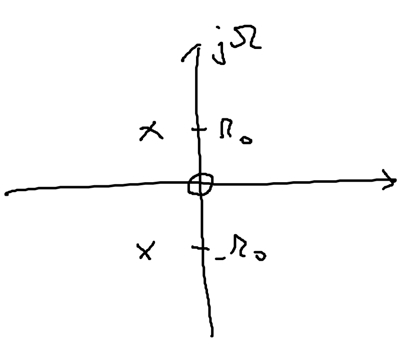


$$j\Omega = \frac{\Omega_c \Omega_c'}{j\Omega'}$$

$$\Omega' = -\frac{\Omega_c \Omega_c'}{\Omega}$$



$$\frac{1}{s + 0.1}$$



$f_0 = 1\ \mathrm{KHZ}$

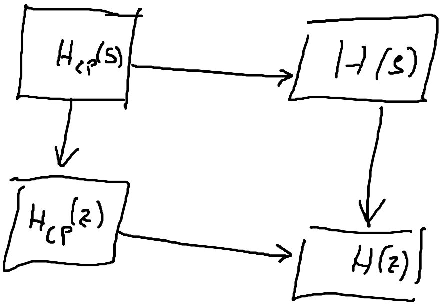
$\Omega_0 = 2\pi f_0$

$$H(s) = \frac{s}{(s-j\Omega_0)(s+j\Omega_0)} = \frac{s}{s^2+\Omega_0^2}$$

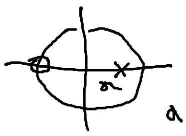
$[1 \quad 0]$

$[1 \quad 0 \quad \Omega_0^2]$

freqs( )   MATLAB



Esempio



$$H(z) = \frac{z+1}{z-a} = \frac{1+z^{-1}}{1-a z^{-1}}$$

$$z = -\frac{z'^2 + c_1 z' + c_2}{c_2 z'^2 + c_1 z' + 1}$$

$$H(\nu) = H(z)\Big|_{z=e^{j2\pi\nu}} = \frac{1+e^{-j2\pi\nu}}{1+a\, e^{-j2\pi\nu}}$$

$$z^{-1} \longrightarrow -\frac{z^{-2} - a_1 z^{-1} + a_2}{a_2 z^{-2} - a_1 z^{-1} + 1}$$

$$H'(z) = \frac{1 - \dfrac{\overbrace{z^{-2} - a_1 z^{-1} + a_2}^{N(z)}}{\underbrace{a_2 z^{-2} - a_1 z^{-1} + 1}_{D(z)}}}{1 + a\dfrac{\overbrace{z^{-2} - a_1 z^{-1} + a_2}^{N(z)}}{\underbrace{a_2 z^{-2} - a_1 z^{-1} + 1}_{D(z)}}}$$

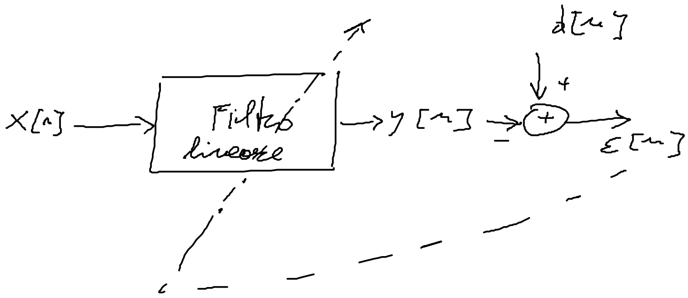
$$= \frac{D(z) - N(z)}{D(z) + a N(z)} =$$

Lezioni 5, 6 e 7

4 Ottobre 14-16 Amedeo
6 Ottobre 9-11 Amedeo
11 Ottobre 14-16 Amedeo

Introduzione alle reti Bayesiane e alla propagazione delle probabilità
Esempi e calcolo dei messaggi

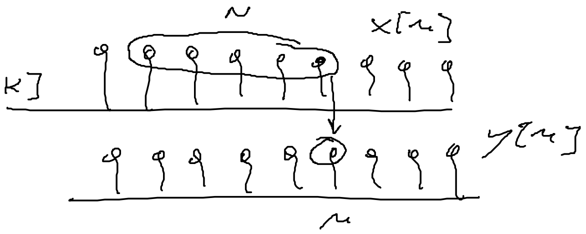
APPRENDIMENTO SUPERVISIONATO o CON TEACHER



$\{ (x[n], d[n]) \quad n = 1, \dots L \}$ insieme di apprendimento (TRAINING SET)

FIR

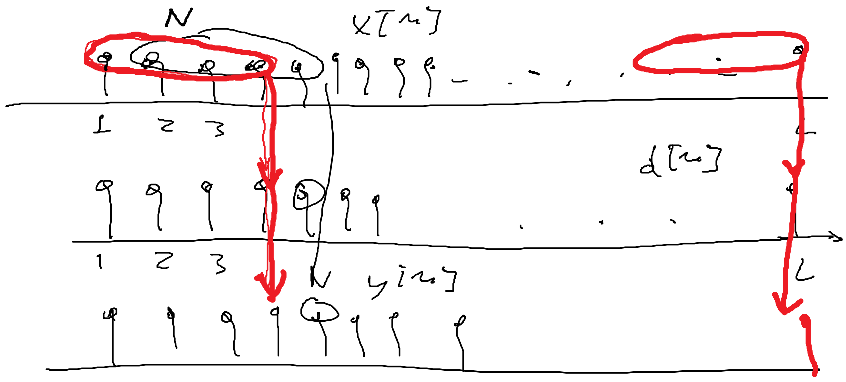
$$y[n] = \sum_{k=0}^{N-1} h[k]\, x[n-k]$$



$$= \left[\quad \underline{h}^T \quad \right] \left[\quad \underline{x}[n] \quad \right]$$

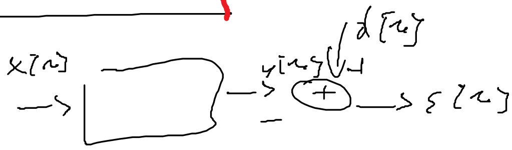
$$= \left[h[0]\, h[1] \dots h[N-1] \right] \begin{bmatrix} x[n] \\ x[n-1] \\ \vdots \\ x[n-N+1] \end{bmatrix} = \underline{h}^T \underline{x}[n]$$

INSIEME DI APPRENDIMENTO



$L \gg N$

$$y[n] = \sum_{k=0}^{N-1} h[k]\, x[n-k]$$



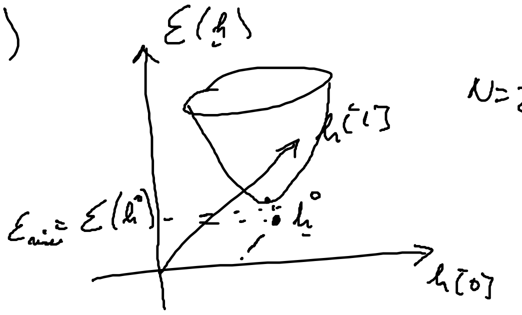
$$\mathcal{E} = \sum_{n=N}^{L} \underbrace{(y[n] - d[n])^2}_{\varepsilon^2[n]}$$

PROBLEMA

$$\{h^o[0] \dots h^o[N-1]\} = \underset{h[0] \dots h[N-1]}{\operatorname{argmin}} \; \mathcal{E}(\underline{h})$$

$$\underline{h}^o = \operatorname{argmin} \; \mathcal{E}(\underline{h})$$

$$\underline{h}^{o} = \arg\min_{\underline{h}} \mathcal{E}[\underline{h}]$$



$N=2$

$$\mathcal{E}(\underline{h}) = \sum_{n=N}^{L} \left(d[n] - \sum_{k=0}^{N-1} h[k]\, x[n-k] \right)^2$$

$$\nabla_{\underline{h}} = \begin{bmatrix} \frac{\partial \mathcal{E}}{\partial h[0]} \\ \vdots \\ \frac{\partial \mathcal{E}}{\partial h[N-1]} \end{bmatrix} = \underline{0} \qquad \frac{\partial \mathcal{E}}{\partial h[\ell]} = \sum_{n=N}^{L} 2 \left(d[n] - \sum_{k=0}^{N-1} h[k]\, x[n-k] \right) \left(-x[n-\ell] \right) = 0$$

$$\ell = 0, \dots N-1$$

$$\sum_{n=N}^{L} \sum_{k=0}^{N-1} h[k]\, x[n-k]\, x[n-\ell] = \sum_{n=N}^{L} d[n]\, x[n-\ell]$$

$$\ell = 0 \dots N-1$$

$$\sum_{k=0}^{N-1} h[k] \underbrace{\sum_{n=N}^{L} x[n-k]\, x[n-\ell]}_{r_x[k,\ell]} = \underbrace{\sum_{n=N}^{L} d[n]\, x[n-\ell]}_{r_{dx}[\ell]} \qquad \ell = 0 \dots N-1$$

$$\sum_{k=0}^{N-1} h^{o}[k]\, r_x[k,\ell] = r_{dx}[\ell] \qquad \ell = 0 \dots N-1$$

$$\underbrace{\begin{bmatrix} r_x[0,0] & \dots & r_x[N-1,0] \\ r_x[0,1] & r_x[1,1] & \dots \\ & \ddots & \\ & & r_x[N-1,N-1] \end{bmatrix}}_{R_x} \begin{bmatrix} h^{o}[0] \\ h^{o}[1] \\ \vdots \\ h^{o}[N-1] \end{bmatrix} = \underbrace{\begin{bmatrix} r_{dx}[0] \\ r_{dx}[1] \\ \vdots \\ r_{dx}[N-1] \end{bmatrix}}_{\underline{r}_{dx}}$$

$$\underline{h}^{o} = R_x^{-1}\, \underline{r}_{dx}$$

$$R_x \geq 0 \qquad \text{D.D.P.O}$$

$$\mathcal{E}(\underline{h}) = \sum_{n=N}^{L} \left(d[n] - \sum_{k=0}^{N-1} h[k]\, x[n-k] \right)^2 =$$

$$= \sum_{n=N}^{L} \Big[d[n]^2 + \underbrace{\Big(\sum_{k=0}^{N-1} h[k]\, x[n-k] \Big)^2}_{(\underline{h}^T \underline{x}[n])^2} - 2\, d[n] \sum_{k=0}^{N-1} h[k]\, x[n-k] \Big]$$

$$= \underbrace{\sum_{n=N}^{L} d[n]^2}_{} + \sum_{n=N}^{L} \underline{h}^T \underline{x}[n]\, \underline{x}^T[n]\, \underline{h} - 2 \sum_{n=N}^{L} d[n]\, \underline{h}^T \underline{x}[n]$$

$$= \mathcal{E}_d + \underline{h}^T \sum_{n=N}^{L} \underline{x}[n]\, \underline{x}^T[n]\, \underline{h} - 2\, \underline{h}^T \sum^{L} d[n]\, \underline{x}[n]$$

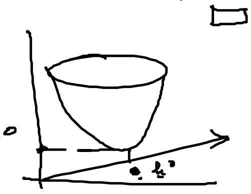
$$= \mathcal{E}_d + \underline{h}^T \underbrace{\sum_{n=N}^{L} \underline{x}[n]\underline{x}^T[n]}_{\underline{\underline{R}}_x} \underline{h} - 2\underline{h}^T \underbrace{\sum_{n=N}^{L} d[n]\underline{x}[n]}_{\underline{r}_{dx}}$$

$$\boxed{\mathcal{E}(\underline{h}) = \mathcal{E}_d + \underline{h}^T \underline{\underline{R}}_x \underline{h} - 2\underline{h}^T \underline{r}_{dx}}$$

$$\nabla_x (x^T A x) = 2\underline{\underline{A}}\,\underline{x}$$

$$\nabla_{\underline{h}} \mathcal{E} = 2\underline{\underline{R}}_x \underline{h}^o - 2\underline{r}_{dx} = 0 \qquad \text{EQ. NORMALE} \qquad \nabla_x x^T v = \underline{v}$$

$$\underline{h}^o = \underline{\underline{R}}_x^{-1} \underline{r}_{dx}$$



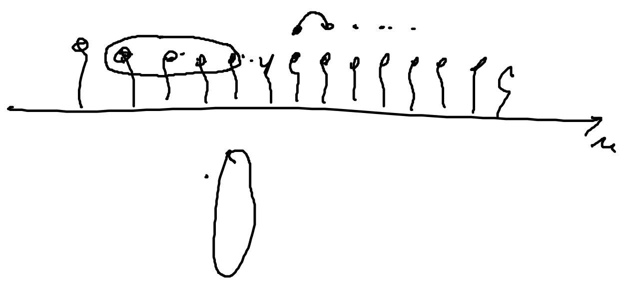
$$H = \underline{\underline{R}}_x > 0 \,? \qquad v^T R_x v > 0 \quad \forall \underline{v}$$

$$v^T \underbrace{\left(\sum_{n=N}^{L} x[n] x^T[n]\right)}_{\underline{\underline{R}}_x} v = \sum_{n=N}^{L} \underline{v}^T \underline{x}[n]\, \underline{x}^T[n]\underline{v} = \sum_{n} (\underline{v}^T \underline{x}[n])^2 > 0$$

$$\mathcal{E}^o = \mathcal{E}_d + r_{dx}^T R_x^{-T} \cancel{R_x} \cancel{R_x^{-1}} \underline{r}_{dx} - 2 r_{dx}^T R_x^{-T} r_{dx}$$

$$\mathcal{E}^o = \mathcal{E}_d - r_{dx}^T R_x^{-1} r_{dx}$$

$$\underline{\underline{R}}_x = \sum_{n=N}^{L} \underline{x}[n]\underline{x}^T[n]$$



$$r_x[k\,\ell] = r_x(k-\ell)$$

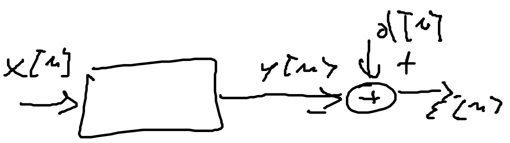
$$\underline{\underline{R}}_x = \begin{bmatrix} r_x(0) & r_x[1] & \cdots & r_x(N-1) \\ r_x(-1) & r_x(0) & \cdots & r_x(N-2) \\ r_x(-2) & r_x(-1) & \ddots & \\ & & & r_x[0] \end{bmatrix} \qquad \text{TOEPLITZ} \qquad \underline{r}_{dx} = \begin{bmatrix} r_{dx}[0] \\ \vdots \\ r_{dx}[N-1] \end{bmatrix}$$

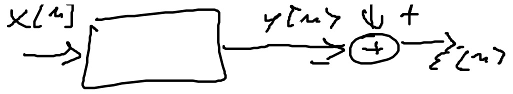
FORMULAZIONE STOCASTICA

$x[n]$ p.a. SSL $\quad E[x[n]x[n-m]] = r_x(m)$

$d[n]$ p.a. SSL congiuntamente a $x[n]$

$$E[d[n]x[n-m]] = r_{dx}(m)$$





$$\mathcal{E}(\underline{h}) = E\left[\varepsilon[n]^2\right] = E\left[(d[n]-y[n])^2\right]$$

$$\underline{h}^o = \underset{\underline{h}}{\operatorname{argmin}}\ \mathcal{E}(\underline{h}) \qquad y[n] = \underline{h}^T \underline{x}[n] = \underline{x}^T[n]\,\underline{h}$$

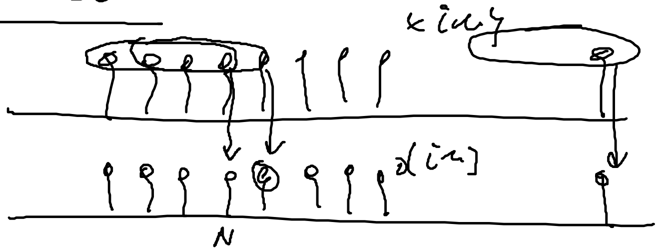
$$\mathcal{E} = E\left[(d[n])^2 + (y[n])^2 - 2\,d[n]\,y[n]\right]$$

$$= \underbrace{E\left[d[n]^2\right]}_{\mathcal{E}_d} + \underline{h}^T \underbrace{E\left[\underline{x}[n]\,\underline{x}^T[n]\right]}_{\underline{R}_x} \underline{h} - 2 \underbrace{E\left[d[n]\,\underline{x}^T[n]\right]}_{\underline{r}_{dx}} \underline{h}$$

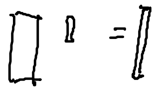
$$= \mathcal{E}_d + \underline{h}^T \underline{R}_x \underline{h} - 2\,\underline{r}_{dx}^T \underline{h}$$

$$\begin{cases} \underline{h}^o = R_x^{-1} r_{dx} \\ \mathcal{E}^o = \mathcal{E}_d - \underline{r}_{dx}^T R_x^{-1} \underline{r}_{dx} \end{cases}$$

APPROCCIO AI MINIMI QUADRATI TOTALMENTE MATRICIALE



$$\underline{A}\,\underline{h} \doteq \underline{d} \qquad \underline{h}^o \qquad \underline{\varepsilon} = \underline{\underline{A}}\,\underline{h} - \underline{d}$$

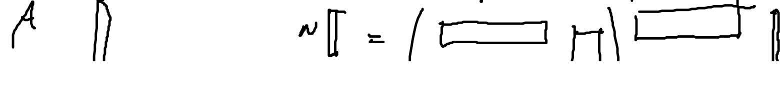


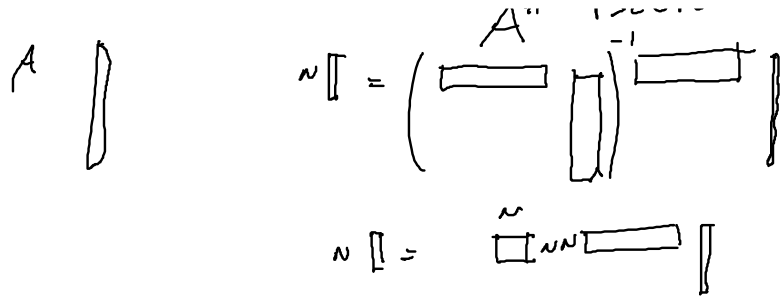
$$\min\ \underline{\varepsilon}^T \underline{\varepsilon}$$

$$\underline{h}^o = \underset{\underline{h}}{\operatorname{argmin}}\ (\underline{A}\,\underline{h} - \underline{d})^T (\underline{A}\,\underline{h} - \underline{d})$$

$$\mathcal{E}(\underline{h}) = \underline{h}^T A^T A\,\underline{h} + d^T d - 2\,d^T A\,\underline{h}$$

$$\nabla_{\underline{h}} = 2\underline{\underline{A}}^T \underline{\underline{A}}\,\underline{h}^o - 2\,\underline{\underline{A}}^T \underline{d} = \underline{0} \Rightarrow \underline{h}^o \doteq \underbrace{(A^T A)^{-1} A^T}_{A^{\#}\ \text{PSEUDO-INVERSA}} d$$





Lezione 9 Filtering

lunedì 18 ottobre 2021 14.06

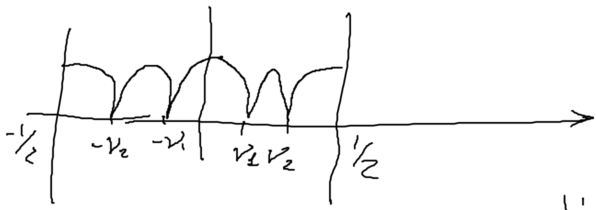
ESERCITAZIONE SUL FILTRAGGIO

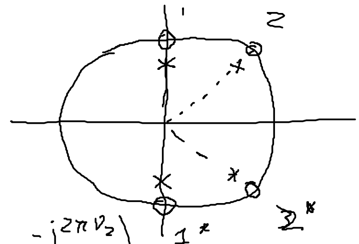
```
clear all
close all
%generate signals and noise
t=1:1/8000:2;
f1=1000;
f2=2000;
s=0.5*cos(2*pi*f1*t+0.3)+0.3*cos(2*pi*f2*t);
subplot(3,1,1)
plot(t, s)
sound(s,8000)
pause
noise=0.1*randn(1,length(t));
subplot(3,1,2)
plot(t, noise)
sound(noise, 5000)
pause
spn=s+noise;
subplot(3,1,3)
plot(t, spn)
sound(spn,8000)
save spn
```

$f_1 = 1\,KHz \quad \nu_1 = \frac{1000}{8000} = \frac{1}{8}$

$f_2 = 2\,KHz \quad \nu_2 = \frac{2000}{8000} = \frac{2}{8}$

FILTRO A
2 NOTCH





$$H(z) = \frac{(z-z_1)(z-z_2)(z-z_1^*)(z-z_2^*)}{(z-\rho_1)(z-\rho_2)(z\,\rho_1^*)(z-\rho_2^*)}$$

$$= \frac{(z-e^{j2\pi\nu_1})(z-e^{-j2\pi\nu_1})(z-e^{+j2\pi\nu_2})(z-e^{-j2\pi\nu_2})}{(z-\rho_1 e^{j2\pi\nu_1})(z-\rho_1 e^{-j2\pi\nu_1})(z-\rho_2 e^{j2\pi\nu_2})(z-\rho_2 e^{-j2\pi\nu_2})}$$

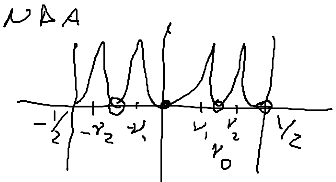
$$= \frac{(z^2 - 2\cos 2\pi\nu_1 z + 1)(z^2 - 2\cos 2\pi\nu_2 z + 1)}{(z^2 - 2\rho_1\cos 2\pi\nu_1 z + \rho_1^2)(z^2 - 2\rho_2\cos 2\pi\nu_2 z + \rho_2^2)}$$

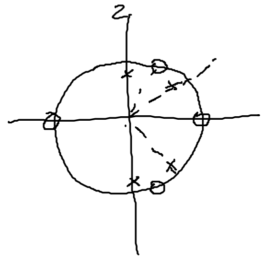
$$= \frac{(1 - 2\cos 2\pi\nu_1 z^{-1} + z^{-2})(1 - 2\cos 2\pi\nu_2 z^{-1} + z^{-2})}{(1 - 2\rho\cos 2\pi\nu_1 z^{-1} + \rho_1^2 z^{-2})(1 - 2\rho\cos 2\pi\nu_2 z^{-1} + \rho_2 z^{-2})}$$

```
close all
clear all
load spn

%design filter
ni1=1/8;
ni2=2/8;
ni0=(ni1+ni2)/2;
ro1=0.99;
ro2=0.99;
B=[conv([1 0 -1],[1 -2*cos(2*pi*ni0) 1])]
% B=[conv([1 -2*cos(2*pi*ni1) 1],[1 -2*cos(2*pi*ni2) 1])]
A=[conv([1 -2*ro1*cos(2*pi*ni1) ro1^2],[1 -2*ro2*cos(2*pi*ni2) ro2^2])]
%verify filter response
[H W]=freqz(B,A,200); %get frequency response (abscissa in rad)
[h t]=impz(B,A);  %get impulse response
figure(1)
subplot(2,2,1)
plot(W/(2*pi),abs(H)/max(abs(H))) %plot magnitude (abscissa in norm freq.)
ylabel('magnitude')
xlabel('nu')
axis([0 0.5 0 1])
grid on
subplot(2,2,3)
plot(W/(2*pi),phase(H)) %plot phase (abscissa in norm freq.)
ylabel('phase')
xlabel('nu')
axis([0 0.5 min(phase(H)) max(phase(H))])
grid on
subplot(2,2,2)
zplane(B,A)      %compute and plot z plane
subplot(2,2,4)
stem(t,h)      %plot impulse response
ylabel('impulse response')
xlabel('n')
% %filter the signal
fs=filter(B,A,spn);
sound(fs,8000)
```

PASSA-BANDA

$\nu_0 = \frac{\nu_1+\nu_2}{2}$





$(z-1)(z+1)$

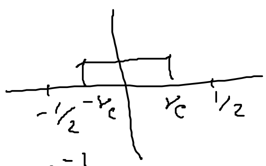
$(z^2 - 2\cos 2\pi\nu_0 z + 1)$

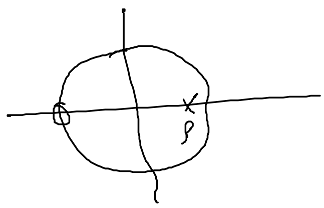
$$H(z) = \frac{(1-z^{2})\,(1-2\cos 2\pi\nu_0 z^{-1} + z^{-2})}{(1-2\rho\cos 2\pi\nu_1 z^{-1}+\rho_1^2 z^{-2})(1-2\rho\cos 2\pi\nu_2 z^{-1}+\rho_2 z^{-2})}$$

Acquire a speech signal

```
%script to create an audio file
Fs=8000; %frequency rate
Nbits=16; %number of bits
recorder=audiorecorder(Fs,Nbits, 1)
disp('Start speaking.')
recordblocking(recorder, 2);
disp('End of Recording.');
x= getaudiodata(recorder); % in x there is my sampled signal
%visualize the signal
plot(x)
%listen to the speech segment
sound(x,Fs)
% %reverse the signal
xr=fliplr(x')
%sound(xr,Fs) %strange!!!!???
figure
%visualize the histogram with 50 bins
hist(x,50)

% %saturate the signal to -1 1
%xs=sign(x)
%
% %listen
%sound(xs,Fs)
%
S=specgram(x); %S is a complex matrix with DFT in the columns
% DFT is computed by default on 256 sample windows
figure
pcolor(abs(S))
```

PASSA-BASSO





$$H(z) = \frac{z+1}{z-\rho} = \frac{1+z^{-1}}{1-\rho z^{-1}}$$

```
close all
clear all
%get signal and add noise
load x

s=x';
noise=0.1*randn(1,length(s));
spn=s+noise
sound(spn,8000)
pause

%progettiamo un passa-basso
ro=0.9;
B=[1 1];
A=[1 -ro];
%verify filter response
[H W]=freqz(B,A,200); %get frequency response (abscissa in rad)
[h t]=impz(B,A);  %get impulse response
figure(1)
subplot(2,2,1)
plot(W/(2*pi),abs(H)/max(abs(H))) %plot magnitude (abscissa in norm freq.)
ylabel('magnitude')
xlabel('nu')
axis([0 0.5 0 1])
grid on
subplot(2,2,3)
plot(W/(2*pi),phase(H)) %plot phase (abscissa in norm freq.)
ylabel('phase')
xlabel('nu')
axis([0 0.5 min(phase(H)) max(phase(H))])
grid on
subplot(2,2,2)
zplane(B,A)      %compute and plot z plane
subplot(2,2,4)
stem(t,h)       %plot impulse response
ylabel('impulse response')
xlabel('n')
%filter the signal
fs=filter(B,A,spn)
sound(fs,8000)
```

Trasf. BILINEARE

$$H_a(s) = \frac{1}{s+a} \qquad H(z) = H_a(s)\Big|_{s = \frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}}$$

$$= \frac{1}{\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} + a} = \frac{T}{2}\,\frac{1}{\frac{1-z^{-1}}{1+z^{-1}} + \frac{Ta}{2}}$$

$$= \frac{T}{2}\,\frac{1}{\frac{2 - 2z^{-1} + aT + aTz^{-1}}{2(1+z^{-1})}} = \frac{T}{\cancel{2}}\,\frac{\cancel{2}(1+z^{-1})}{(2+Ta) + (Ta-2)z^{-1}}$$

$$= \frac{T_a + T_a z^{-1}}{(2+T_a) + (T_a - 2)z^{-1}}$$

```
close all
clear all
%filtering2
%get signal
load x

s=x';
noise=0.1*randn(1,length(s));
spn=s+noise
sound(spn,8000)
pause

%progettiamo un passa-basso con la trsformazione bilineare
T=1/8000;
a=1000;
B=[T T];
A=[2+T*a  T*a-2]
%verify filter response
[H W]=freqz(B,A,200); %get frequency response (abscissa in rad)
[h t]=impz(B,A);  %get impulse response
figure(1)
subplot(2,2,1)
plot(W/(2*pi),abs(H)/max(abs(H))) %plot magnitude (abscissa in norm freq.)
ylabel('magnitude')
xlabel('nu')
axis([0 0.5 0 1])
grid on
subplot(2,2,3)
plot(W/(2*pi),phase(H)) %plot phase (abscissa in norm freq.)
ylabel('phase')
xlabel('nu')
axis([0 0.5 min(phase(H)) max(phase(H))])
grid on
subplot(2,2,2)
zplane(B,A)      %compute and plot z plane
subplot(2,2,4)
stem(t,h)       %plot impulse response
ylabel('impulse response')
xlabel('n')
%filter the signal
fs=filter(B,A,spn)
sound(fs,8000)
```

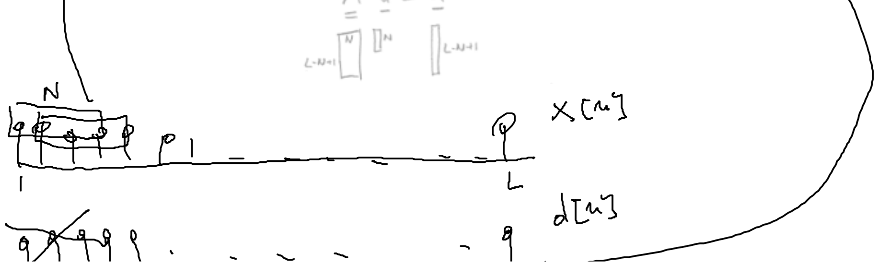
ORGANIZE DATA IN A AND b

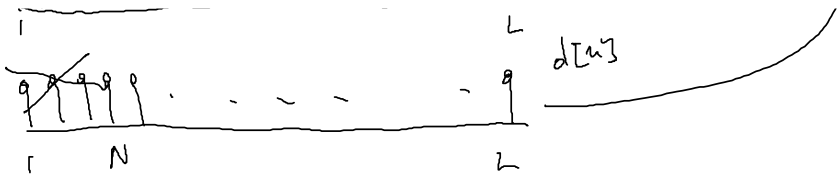
WIENER FILER WITH THE PSEUDO INVERSE

$$\begin{bmatrix} x[N] & x[N-1] & \dots & x[1] \\ x[N+1] & x[N] & \dots & x[2] \\ x[N+2] & x[N+1] & \dots & x[3] \\ x[N+3] & x[N+2] & \dots & x[4] \\ \vdots & & & \\ x[L-1] & x[L-2] & \dots & x[L-N] \\ x[L] & x[L-1] & \dots & x[L-N+1] \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[N] \end{bmatrix} \doteq \begin{bmatrix} d[N] \\ d[N+1] \\ d[N+2] \\ \vdots \\ d[L] \end{bmatrix} \; L-N+1$$

$$\underline{\underline{A}}\;\underline{h} \doteq \underline{d}$$

L-N+1 N N L-N+1





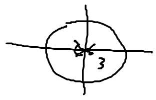

```
s=[1 2 3 4 1 2 3 4 1 3 4 ]
N=4;
d=rand(length(s)-N+1,1)
R=fliplr(s(1,1:N));
C=s(1,N:length(s));
A=toeplitz(C,R)

h=pinv(A)*d
```

```
s =

   1   2   3   4   1   2   3   4   1   3   4

d =

   0.8147
   0.9058
   0.1270
   0.9134
   0.6324
   0.0975
   0.2785
   0.5469


A =

   4   3   2   1
   1   4   3   2
   2   1   4   3
   3   2   1   4
   4   3   2   1
   1   4   3   2
   3   1   4   3
   4   3   1   4


h =

   0.0957
   0.1270
  -0.0577
   0.0321
```

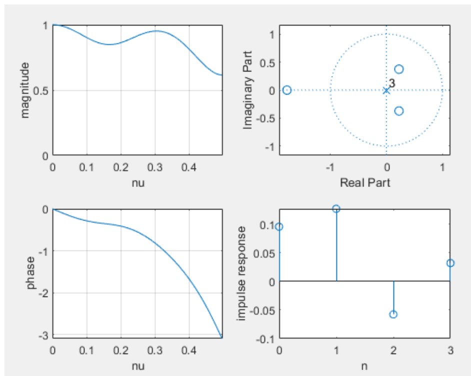
PREPARE THE ALGORITHM ON A SMALL SET

$$H(z) = h[1] + h[2]z^{-1} + h[3]z^{-2} + h[4]z^{-3}$$

$$= \frac{1}{z^3}\left(h[1]z^{3} + h[2]z^{2} + h[3]z + h[4]\right)$$



Questo filtro non ha alcun senso perché s e d sono inventati.



Lezione 10 Filtering

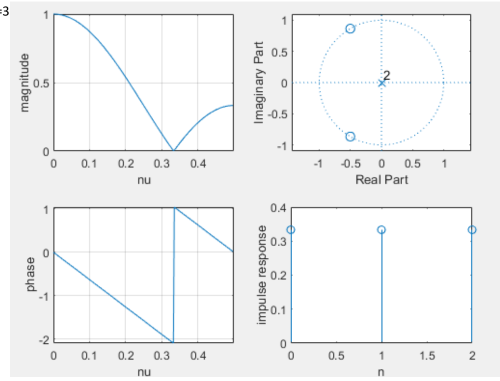
mercoledì 20 ottobre 2021 08.58

Prepare a function for filter characterization

```
function FilterChar(B,A)
[H W]=freqz(B,A,200); %get frequency response (abscissa in rad)
[h t]=impz(B,A);  %get impulse response
figure
subplot(2,2,1)
plot(W/(2*pi),abs(H)/max(abs(H))) %plot magnitude (abscissa in norm freq.)
ylabel('magnitude')
xlabel('nu')
axis([0 0.5 0 1])
grid on
subplot(2,2,3)
plot(W/(2*pi),phase(H)) %plot phase (abscissa in norm freq.)
ylabel('phase')
xlabel('nu')
axis([0 0.5 min(phase(H)) max(phase(H))])
grid on
subplot(2,2,2)
zplane(B,A)      %compute and plot z plane
subplot(2,2,4)
stem(t,h)       %plot impulse response
ylabel('impulse response')
xlabel('n')
```

Test : Moving average filter N=3

```
A=[1]
B=[1/3 1/3 1/3]
FilterChar(B,A)
```

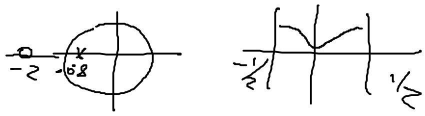


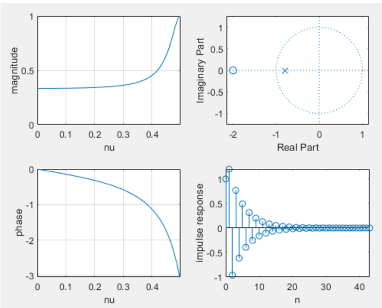
$$H(z) = \frac{1}{3}(1 + z^{-1} + z^{-2}) = \frac{1}{3}\frac{(z^2+z+1)}{z^2}$$

Test on recursive filter:

```
A=[1 0.8]
B=[1 2]
```

$$H(z) = \frac{1+2z^{-1}}{1+0.8z^{-1}} = \frac{z+2}{z+0.8}$$

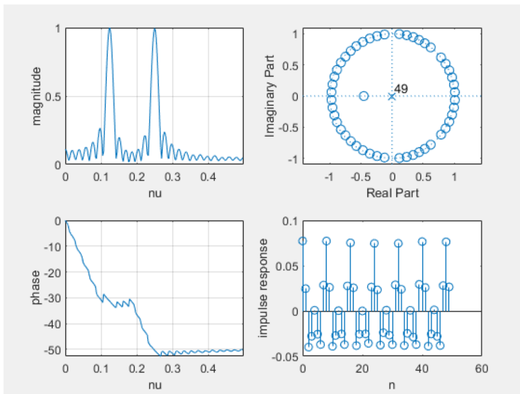




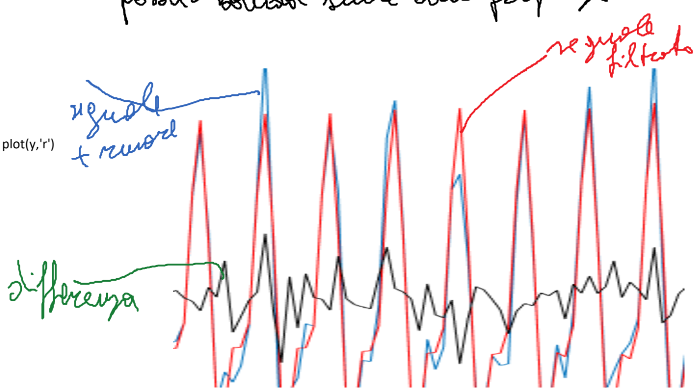
FILTRO DI WIENER (AI MINIMI QUADRATI) USANDO LA PSEUDO-INVERSA

```
clear all
close all
%generate a noisy signal
t=1:1/8000:2;
f1=1000;
f2=2000;
s=0.5*cos(2*pi*f1*t+0.3)+0.3*cos(2*pi*f2*t);
subplot(3,1,1)
plot(t, s)
sound(s,8000)
pause
noise=0.1*randn(1,length(t));
subplot(3,1,2)
plot(t, noise)
sound(noise, 5000)
pause
x=s+noise;
subplot(3,1,3)
plot(t, x)
sound(x,8000)
save s
```

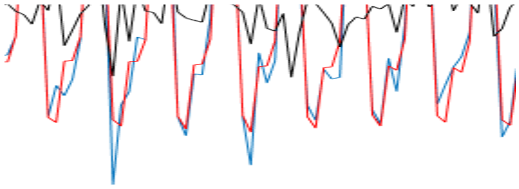
```
close all
clear all
%wiener using the pseudo-inverse
%import the clean signal
load s
%load x
%s=x(4000:8000)';
sound(s,8000)
pause

N=50;
d=s(1, N:length(s)); %desired vector
%contaminate the signal with noise
spn=s+0.1*randn(1,length(s));
sound(spn, 8000)
pause

R=fliplr(spn(1,1:N));
C=spn(1,N:end);
A=toeplitz(C,R);
h=pinv(A)*d'

FilterChar(h',[1])
y=filter(h',[1],spn);
% figure
% plot(spn)
% hold on
% plot(y, 'r')
sound(spn)
pause
sound(y)

figure, plot(spn), hold on , plot(y,'r')
hold on, plot(spn-y,'k')
```



Qui si vede come il filtro seleziona le due sinusoidi diventando un passa-banda sulle due frequenze

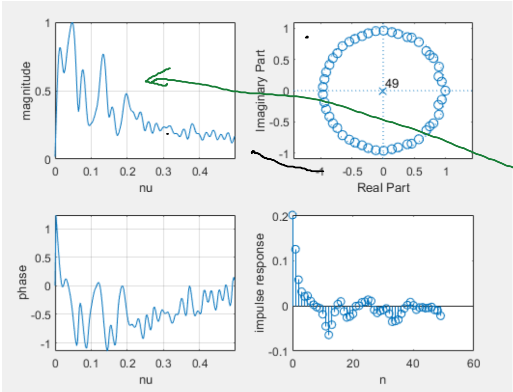


Secondo modo in ...

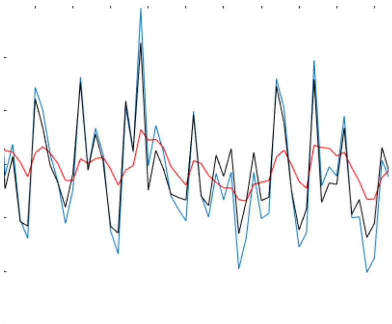


Segnale vocale in x

```
close all
clear all
%wiener using the pseudo-inverse
%import the clean signal
%load s
load x
s=x(4000:8000)'; %use only a voiced segment
sound(s,8000)
pause

N=50;
d=s(1, N:length(s)); %desired vector
%contaminate the signal with noise
spn=s+0.1*randn(1,length(s));
sound(spn, 8000)
pause

R=fliplr(spn(1,1:N));
C=spn(1,N:end);
A=toeplitz(C,R);
h=pinv(A)*d'

FilterChar(h',[1])
y=filter(h',[1],spn);
% figure
% plot(spn)
% hold on
% plot(y, 'r')
sound(spn)
pause
sound(y)

figure, plot(spn), hold on , plot(y,'r')
hold on, plot(spn-y,'k')
```



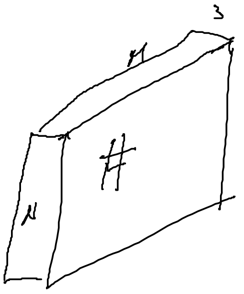
il filtro si focalizza sulle "formanti"



Risultato non molto buono soprattutto all'ascolto poiché ci sono effetti percettivi non lineari.

Lezione 11

lunedì 25 ottobre 2021 14.22

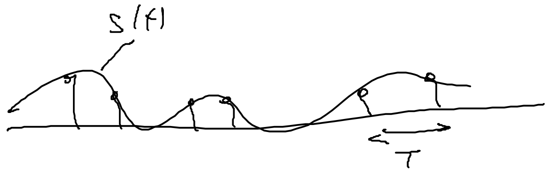


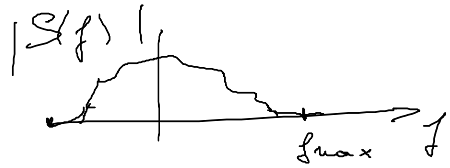
$I[n,m,p]$

$n = 1, \ldots N$

$m = 1, \ldots M$

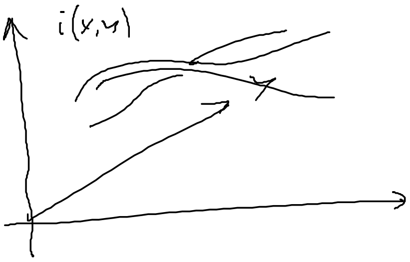
$p = 1, 2, 3$





$$\frac{1}{T} > 2 f_{max}$$

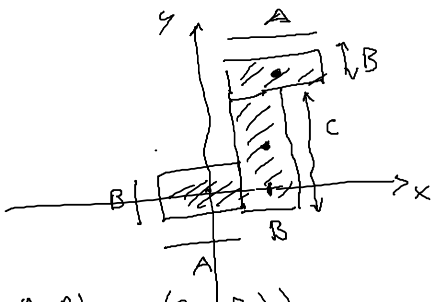
$$S(f) = \int_{-\infty}^{+\infty} s(t)\, e^{-j2\pi f t}\, dt$$



$$\mathcal{F}^2\left[i(x,y)\right] = I(u,v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} i(x,y)\, e^{-j2\pi(ux+vy)}\, dx\, dy$$

freq. orizzontali — freq. verticale

$$\mathcal{F}_2\left[\;\right] =$$

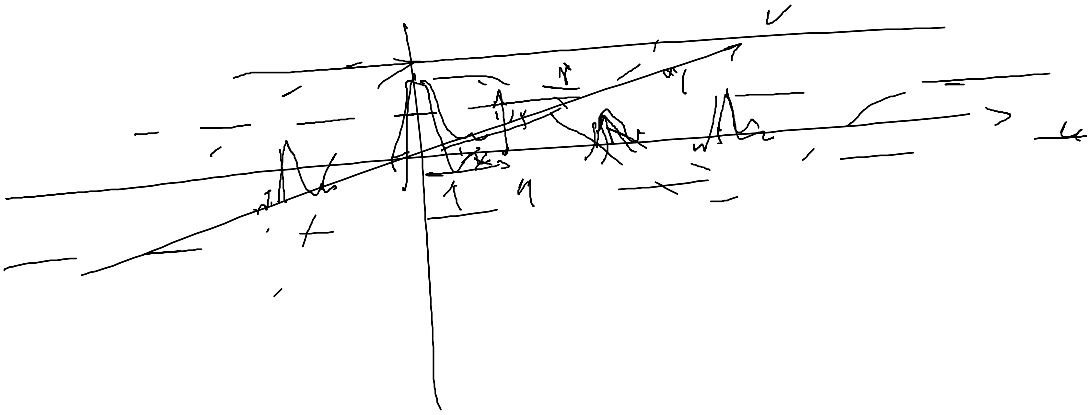


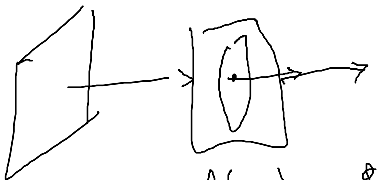
$$f(x,y) = \mathrm{rect}\left(\frac{x}{A}, \frac{y}{B}\right) + \mathrm{rect}\left(\frac{x-\left(\frac{A}{2}+\frac{B}{2}\right)}{B}, \frac{y-\left(\frac{C}{2}-\frac{B}{2}\right)}{C}\right) + \mathrm{rect}\left(\frac{x-A}{A}, \frac{y-C}{B}\right)$$

$$F(u,v) = A\,\mathrm{sinc}\,Au\; B\,\mathrm{sinc}\,Bv + B\,\mathrm{sinc}\,Bu\; C\,\mathrm{sinc}\,Cv\; e^{-j2\pi\left(\frac{A}{2}+\frac{B}{2}\right)u}\, e^{-j2\pi\left(\frac{C}{2}-\frac{B}{2}\right)v} + A\,\mathrm{sinc}\,Au\; B\,\mathrm{sinc}\,Bv\; e^{-j2\pi A u}\, e^{-j2\pi C v}$$

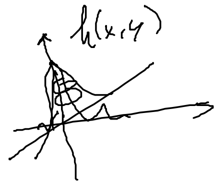
$$\left(\mathrm{rect}\left(\frac{x}{A}, \frac{y}{B}\right) * \mathrm{comb}_{XY}(x,y)\right) \cdot \mathrm{rect}\left(\frac{x}{C}, \frac{y}{D}\right)$$

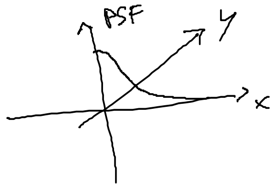
$$\left(\frac{1}{XY}\,\mathrm{comb}_{\frac{1}{X}\frac{1}{Y}}(u,v)\cdot AB\,\mathrm{sinc}\,Au\;\mathrm{sinc}\,Bv\right) * CD\,\mathrm{sinc}\,Cu\;\mathrm{sinc}\,Dv$$

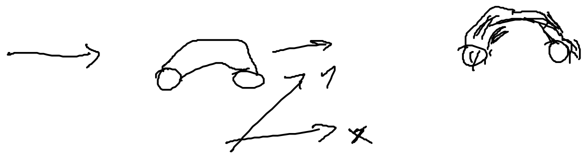




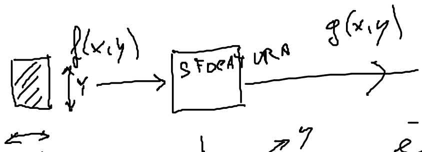
$f(x,y)$ $\qquad h(x,y)$ $\qquad g(x,y) = (h \underset{2D}{*} f)(x,y)$



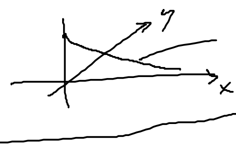




Esempio sfocatura da "mosso a destra"



$$g(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(\xi,\eta)\, h(\xi - x, \eta - y)\, d\xi\, d\eta$$

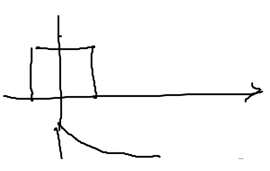


$$e^{-\alpha x} = h(x,y)$$

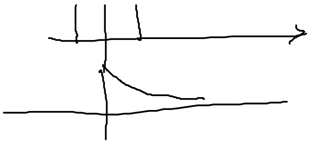
$$= \iint \Pi\left(\frac{\xi}{X}\right)\Pi\left(\frac{\eta}{Y}\right) e^{-\alpha(\xi - x)}\, u(\xi - x)\, \delta(\eta - y)\, d\xi\, d\eta$$

$$= \int_{-\infty}^{+\infty} \Pi\left(\frac{\xi}{X}\right) e^{-\alpha(\xi - x)}\, u(\xi - x)\, d\xi \int_{-\infty}^{+\infty} \Pi\left(\frac{\eta}{Y}\right)\delta(\eta - y)\, d\eta$$

$$\Pi\left(\frac{y}{Y}\right)$$

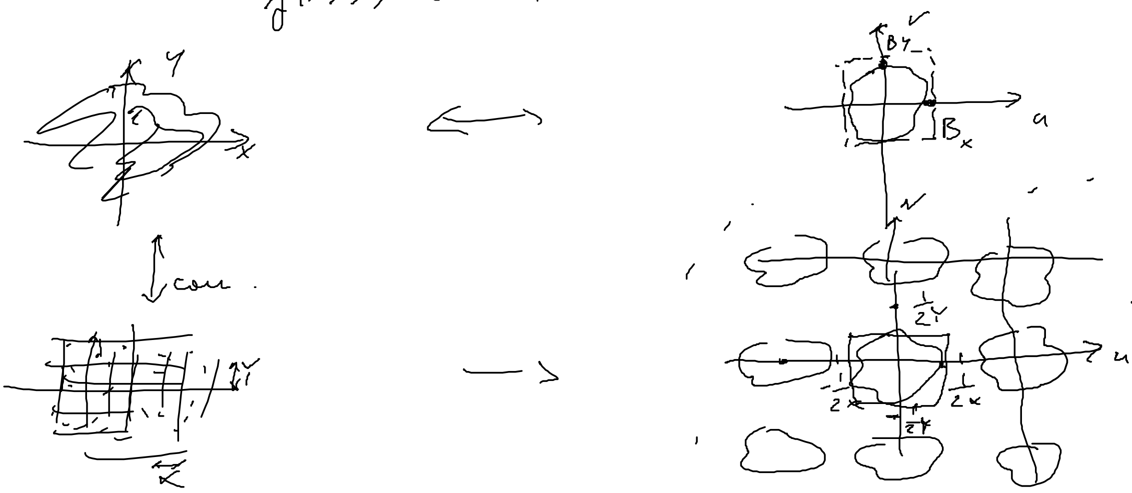


$$g(x,y) = b(x)\,\Pi\left(\frac{y}{Y}\right)$$



$$g(x,y) = b(x)\,\Pi\left(\frac{y}{Y}\right)$$

$$f(x,y) \overset{\mathcal{F}_2}{\longleftrightarrow} F(u,v)$$



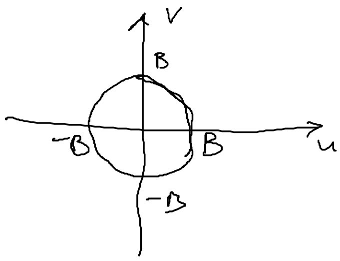
$$\begin{cases} B_x < \frac{1}{2X} \\ B_y < \frac{1}{2Y} \end{cases}$$

$X = Y$

$$\frac{1}{X} > 2B_x$$

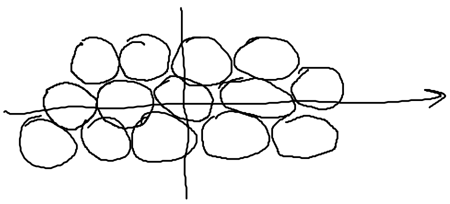
$$\frac{1}{\cancel{X}} > 2B_y$$

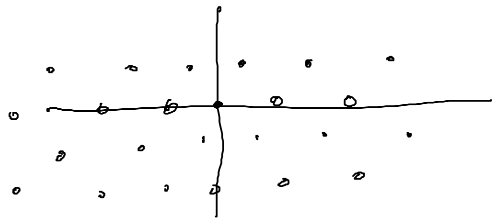
$$\begin{cases} \frac{1}{X} > 2B_x \\ \frac{1}{Y} > 2B_y \end{cases}$$

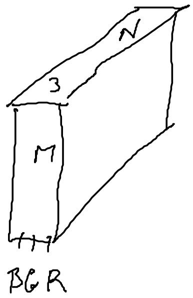


$F(u,v)$

$$\frac{1}{X} \geq 2B$$







$$L = c_1 R + c_2 G + c_3 B$$
$$X = a R + x G + \cdot B$$
$$Y = \quad - \quad -$$

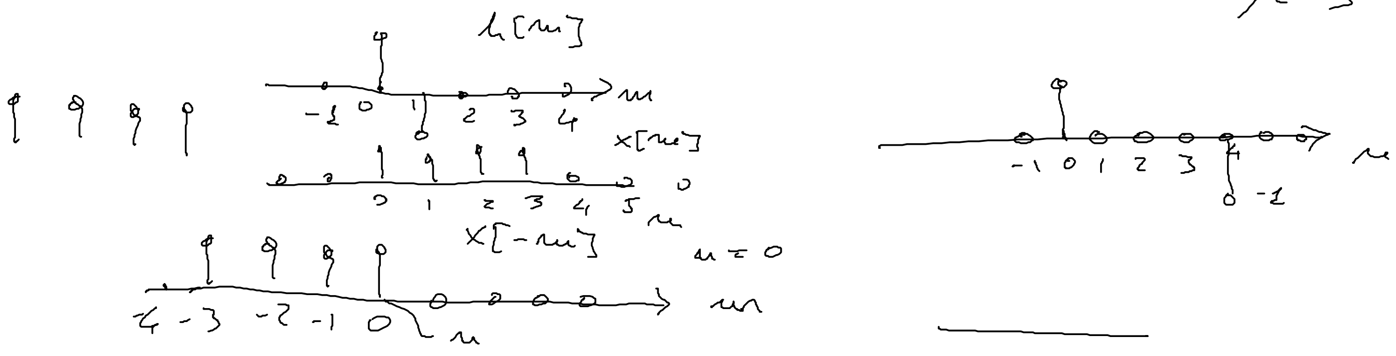
$$\begin{bmatrix} L \\ X \\ Y \end{bmatrix} = \begin{bmatrix} & T & \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Esercitazione in Matlab.
da dispensa 1 $\Rightarrow$ Colori e modifiche
Contrasto e modifiche

da dispensa 2 $\Rightarrow$

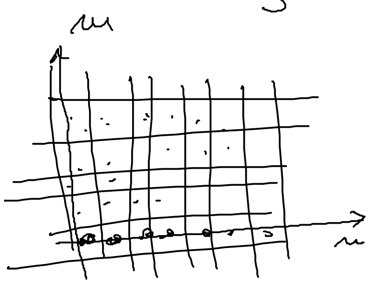
$$y[n] = (h * x)[n] = \sum_{m=-\infty}^{+\infty} h[m]\, x[n-m] = \sum_{m=-\infty}^{+\infty} x[m]\, h[n-m]$$
$$x[-(m-n)]$$

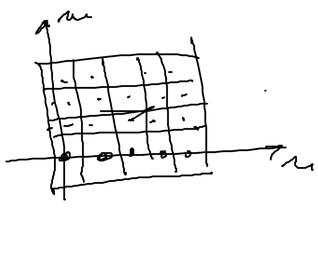
Esempio

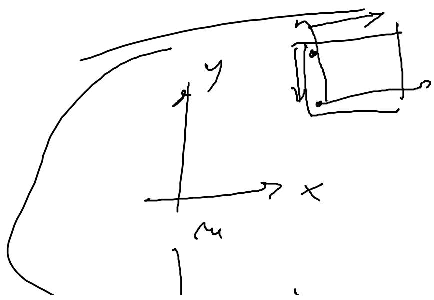


$$x[n, m] \qquad h[n, m]$$

$$y[n,m] = (h * x)[n,m] = \sum_{\ell=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} h[\ell, q]\, x[n-\ell, m-q]$$
$$= \sum_{\ell} \sum_{q} x[\ell, q]\, h[n-\ell, m-q]$$







Esempi in Matlab …….

Esempi in Matlab …….

$\longrightarrow_t \; X(t)$ p.a. S.S.L. $E[X(t)]=0$

$$\longrightarrow_t \; X(f) = \int_{-\infty}^{+\infty} x(t)\, e^{-j2\pi f t}\, dt$$

$E[X(f)] = 0$

$$E[X(f)X^*(f-\Delta)] = E\left[\int_{-\infty}^{+\infty} x(t_1) e^{-j2\pi f t_1} dt_1 \int_{-\infty}^{+\infty} x^*(t_2) e^{j2\pi(f-\Delta)t_2} dt_2\right]$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} E[x(t_1)x^*(t_2)]\, e^{-j2\pi f t_1 + j2\pi(f-\Delta)t_2}\, dt_1\, dt_2$$

$\tau = t_1 - t_2 \qquad t = t_2$

$$= \iint E[x(t)x^*(t-\tau)]\, e^{-j2\pi f \tau}\, e^{-j2\pi\Delta t}\, d\tau\, dt$$

$$\int R_x(\tau) e^{-j2\pi f \tau}\, d\tau \int e^{-j2\pi\Delta t}\, dt$$

$$= P_x(f)\, \delta(\Delta) \quad \text{---} \quad = 0 \quad \forall \Delta \neq 0$$

DFT
$$X(\nu) = \sum_{n=-\infty}^{+\infty} x[n]\, e^{-j2\pi n \nu} \qquad x[n] = \int_{-1/2}^{1/2} X(\nu)\, e^{j2\pi\nu n}\, d\nu$$

$x[n]$ p.a. SSL $E[X[n]] = 0$

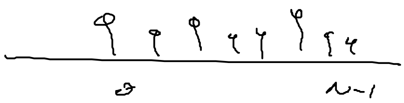
$R_x[m] = E[X[n]\, x^*[n-m]]$

$E[X(\nu)X^*(\nu-\Delta)] = \; .\;.\;.\;.$ dagli appunti....

$$= P_x(\nu) \sum_{k=-\infty}^{+\infty} \delta(k-\Delta)$$

! ! !

$$X_N(\nu) = \sum_{n=0}^{N-1} x[n]\, e^{-j2\pi\nu n}$$

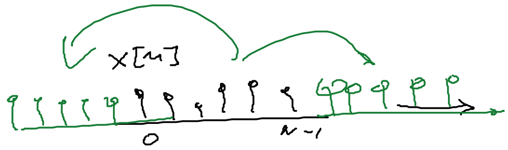


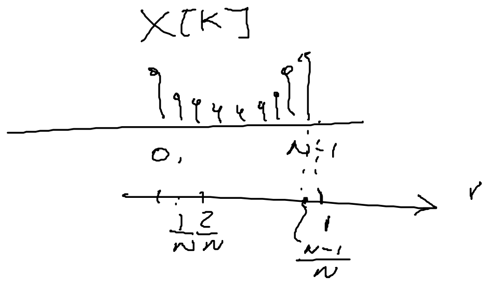
$E[X_N(\nu)X_N^*(\nu)] = \; .\;.\;.$ dagli appunti.....

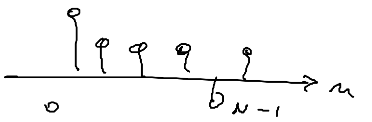
DFT . . . . . dagli appunti . . . . .

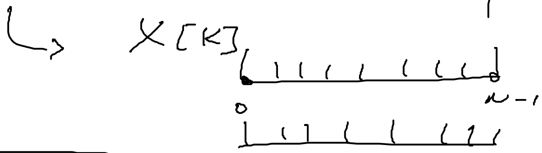
$X[K]$

$x[n]$







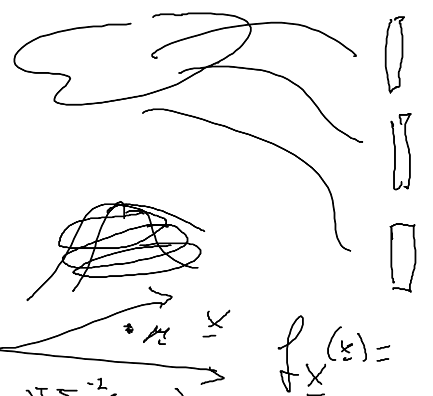


Qualche esempio di DFT su piccole sequenze in MATLAB (periodicità, zero-padding, etc ...)

$\underline{X} =$

Richiamo

$\underline{x}$ vettore aleatorio



$$\underline{\mu} = E[\underline{x}] = \begin{bmatrix} E[x_1] \\ E[x_2] \\ \vdots \\ E[x_N] \end{bmatrix}$$

$$\underline{\underline{\Sigma}} = \mathrm{cov}[\underline{x}] = E[(\underline{x}-\underline{\mu})(\underline{x}-\underline{\mu})^T]$$

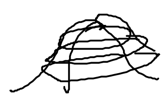
$f_{\underline{x}}(\underline{x}) =$

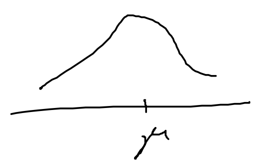
$$f_{\underline{x}}(\underline{x}) = \frac{1}{(2\pi)^{N/2}|\underline{\underline{\Sigma}}|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \underline{\underline{\Sigma}}^{-1}(\underline{x}-\underline{\mu})}$$

$$\frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\mu)^2}{2\sigma_x^2}}$$

$$\underline{\underline{R}} = E[\underline{x}\,\underline{x}^T] = \underline{\underline{\Sigma}} \text{ se } \underline{\mu} = \underline{0}$$

H





$$\underline{\underline{\Sigma}} = \underline{\underline{R}} - \underline{\mu}\,\underline{\mu}^T \qquad \sigma_x^2 = E[x^2] - \mu^2$$

$$\underline{x} \to \boxed{W_N} \xrightarrow{\underline{X}} \boxed{\frac{1}{N}W_N^H} \to \underline{x}$$

$$\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \qquad \begin{cases} a_{11} x_1 = y_1 \\ a_{21} x_1 + a_{22} x_2 = y_2 \\ a_{31} x_1 + a_{32} x_2 + a_{33} x_3 = y_3 \end{cases} \qquad x_1 = \frac{y_1}{a_{11}}$$

$$\begin{cases} x_1 = \dfrac{y_1}{a_{11}} \\ x_2 = y_2 - a_{21} x_1 \\ x_3 = y_3 - a_{31} x_1 - a_{32} x_2 \end{cases}$$

$$A \underline{x} = \underline{y}$$

$$L L^T x = \underline{y}$$

$$L^T x = L^{-1} y$$

Lezione 14

lunedì 8 novembre 2021 14.08

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \qquad \underline{x} \to \boxed{C} \to \underline{y}$$

Fatt. di $\underline{\underline{R}}_x$

$$\underline{\underline{R}}_x = E[\underline{x}\,\underline{x}^T]$$

$$y = Cx$$

$$\underline{\underline{R}}_y = \text{diagonale}$$

$$\underline{\underline{R}}_x = \underline{\underline{F}}\,\underline{\underline{F}}^T$$

$$y = F^{-1}x$$

$$E[yy^T] = E[F^{-1}xx^TF^{-T}] = F^{-1}\underbrace{E[xx^T]}_{\underline{\underline{R}}_x}F^{-T} = \cancel{F^{-1}F}\,\cancel{F^TF^{-T}} = I$$

$$C = \alpha F^{-1} \Rightarrow E[yy^T] = \alpha^2 I$$

$$\Lambda^{1/2}\Lambda^{1/2}$$

$$C = F^{-1}$$

$$\underline{\underline{R}}_x = F\Lambda F^T$$

$$E[yy^T] = F^{-1}R_xF^{-T} = \cancel{F^{-1}}F\Lambda \cancel{F^T}\cancel{F^{-T}} = \Lambda$$

$$C = \Lambda^{-1/2}F^{-1}$$

$$E[yy^T] = E[\Lambda^{-1/2}F^{-1}xx^TF^{-T}\Lambda^{-1/2}] = \Lambda^{-1/2}\cancel{F^{-1}}F\Lambda\cancel{F^T}\cancel{F^{-T}}\Lambda^{-1/2} = \underline{\underline{I}}$$

$$\underline{\underline{R}}_x = F\Lambda F^T \qquad \text{Quale fattorizzazione}$$

FATT. DI CHOLESKY

$$\underline{\underline{R}}_x = L\Lambda L^T \qquad C = L^{-1} \qquad C = \Lambda^{-1/2}L^{-1}$$

$$\underline{x} \to \boxed{C} \to \underline{y}$$

FATT. SPETRALE

$$\underline{\underline{R}}_x = Q\Lambda Q^T \qquad C = Q^{-1} = Q^T \qquad C = \Lambda^{-1/2}Q^T$$

$$Q^TQ = I$$

$$\underline{\underline{\Lambda}} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix}$$

$$\lambda_1 > \lambda_2 > \dots > \lambda_N$$

$$\underline{x} \to \boxed{Q^T} \to \underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \} \text{piccole}$$

$$\underline{y} = Q^T\underline{x} \qquad E[\underline{y}\,\underline{y}^T] = \underline{\underline{\Lambda}} = \begin{pmatrix} E[y_1^2] & & 0 \\ & E[y_2^2] & \end{pmatrix}$$

$$\underline{y} = Q^T \underline{x} \qquad E[\underline{y}\underline{y}^T] = \underset{=}{\Lambda} = \begin{pmatrix} E[y_1^2] & & 0 \\ & E[y_2^2] & \\ 0 & & \ddots & \\ & & & E[y_N^2] \end{pmatrix}$$

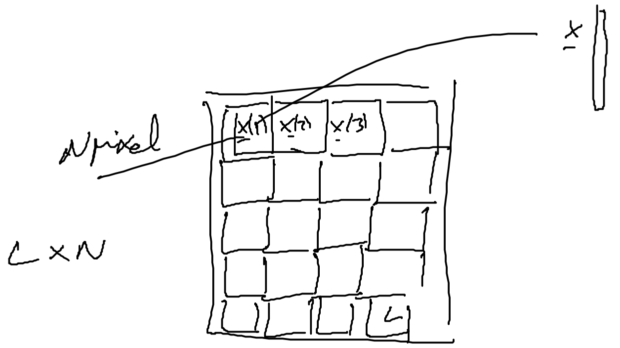
conservo solo M componenti (quelle a varianza più grande)
(autovalore più grande)

$$\underline{x} \rightarrow \boxed{Q_M^T} \rightarrow y_M \rightarrow \boxed{Q_M} \rightarrow \underline{x}_r$$

COMPONENTI PRINCIPALI DI $\underline{x}$

TRASFORMATA DI KARHUNEN-LOÈVE (KLT)

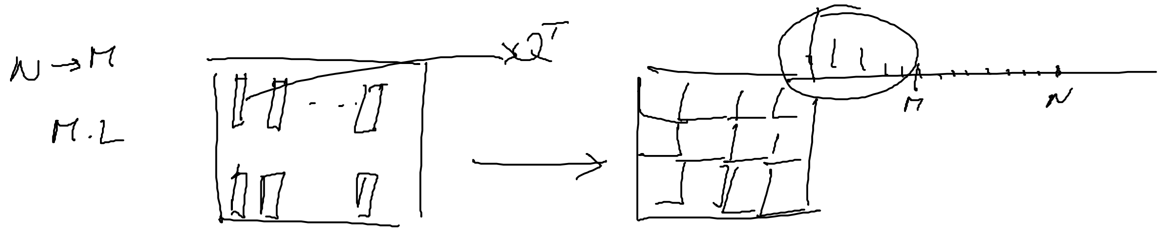
$$\mathcal{E}_M = E\left[(\underline{x} - \underline{x}_r)^T (\underline{x} - \underline{x}_r)\right]$$

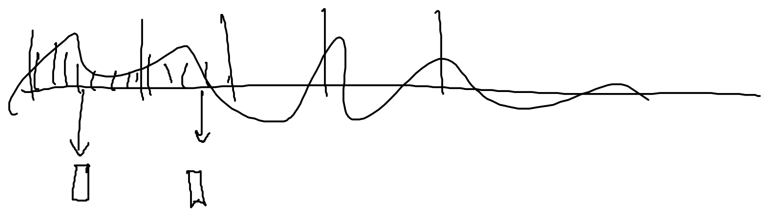


$$\underline{R}_x = \frac{1}{L}\sum_{i=1}^{L} \underline{x}(i)\,\underline{x}^T(i)$$

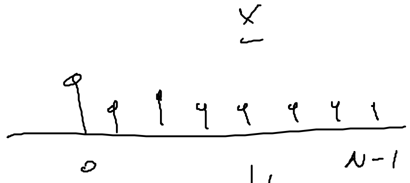
$$\underline{R}_x = Q \Lambda Q^T$$

$$\lambda_1 > \lambda_2 > \lambda_3 \dots > \lambda_N$$



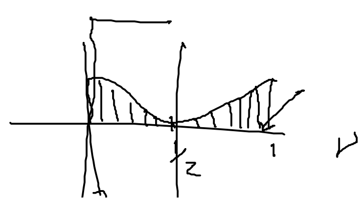


Altre trasformate



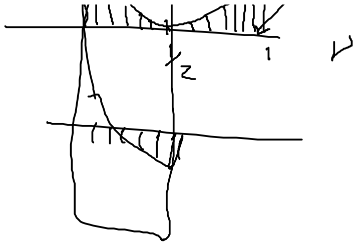
$$\Downarrow \text{ DFT (FFT)}$$

$$\underline{X} = \underset{=}{C}\,\underline{x}$$



$0 \quad \ldots \quad N-1$

$0 \quad \frac{1}{N} \quad \ldots \quad \frac{N-1}{N}$



$$\boxed{\underline{X} = \begin{bmatrix} +1 & -1 & \\ & & \end{bmatrix} \underline{x}}$$

$$\begin{bmatrix} \cdot \end{bmatrix} = \begin{bmatrix} 1 & -1 & \cdots \\ \vdots & & \\ & \bullet & \end{bmatrix} \begin{bmatrix} \underline{x} \end{bmatrix}$$

autovettori di $\underline{\underline{R}}_x$

KLT

DFT

DCT

DST

DHT

DST

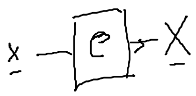
⋮

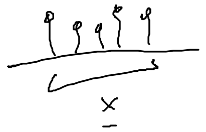
Dalla dispensa: Trasformate applicate alle immegini
Simulazioni in Matlab con la DCT 1D e 2D
Compressione

DFT
DCT
DST
DHT
⋮

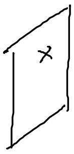
$$X(\nu) = \sum_{n=0}^{N-1} x[n]\, e^{-j2\pi\nu n} \qquad \text{DTFT}$$

$$X[k] = \sum_{n=0}^{N-1} x[n]\, e^{-j2\pi\frac{kn}{N}} \qquad \text{DFT}$$





$$\begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix} \longrightarrow \begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix}$$



$$\{x[i,j]\}_{i=0:N-1}^{j=0:M-1} \qquad x[\cdot$$

$$\begin{bmatrix} x[0,0] & x[0,1] & \dots & x[0,M-1] \\ x[1,0] & x[1,1] & \dots & x[1,M-1] \\ \vdots & & & \\ x[N-1,0] & \dots & \dots & x[N-1,M-1] \end{bmatrix}$$

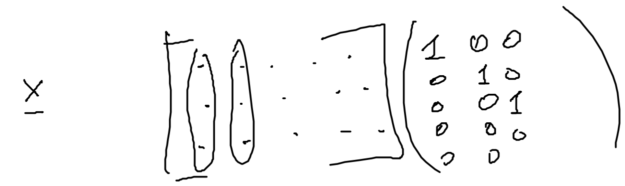
$$X[k,l] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[n,m]\, \underbrace{e^{-j\frac{2\pi kn}{N}}}_{W_N^{kn}}\, \underbrace{e^{-j\frac{2\pi lm}{M}}}_{W_M^{lm}}$$

DFT2

DTFT2

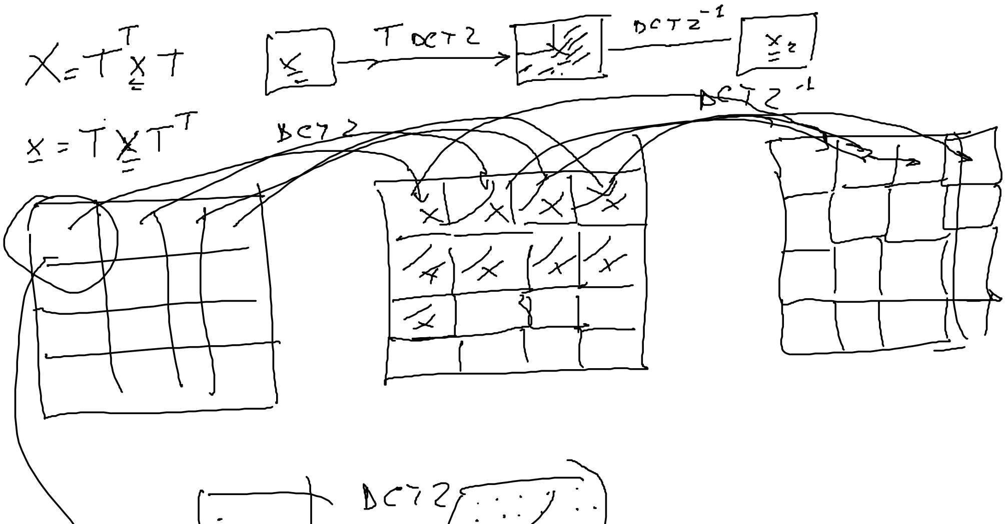
$$X(\nu,\eta) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[n,m]\, \underbrace{e^{-j2\pi\nu n}\, e^{-j2\pi\eta m}}_{W}$$

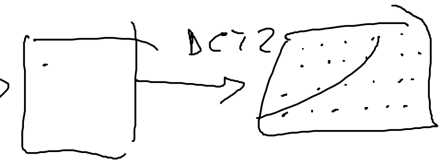
freq. verticale, freq. orizzontale



$$X = T^T x T$$

$$x = T X T^T$$



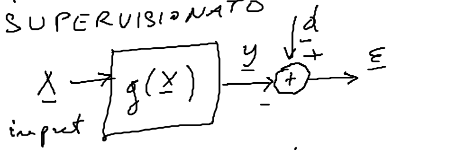


500 Kbyte.

$N \times M \times 3$ $500 \times 1000 \times 3$

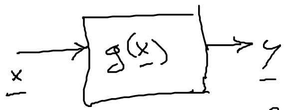
500'000 1'500'000 MByte

PROBLEMA SUPERVISIONATO



Problema dipende da $\underline{x}$ e $\underline{d}$

PROBLEMA NON SUPERVISIONATO



Dipende solo da $\underline{x}$

$\underline{x} \in \mathbb{R}^N$ $\quad \underline{d} \in \mathbb{R}^n$ $\quad$ REGRESSIONE

$\underline{x} \in \mathbb{R}^N$ $\quad \underline{d} \in \mathcal{A} = \{a_1 \ldots a_M\}$ $\quad$ CLASSIFICAZIONE

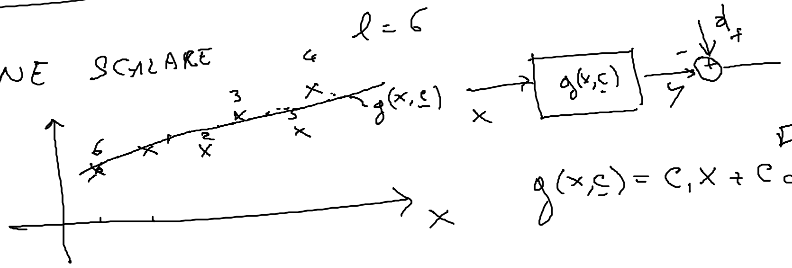
$\begin{cases} \text{DAI DATI} & \{\underline{x}[n], \underline{d}[n]), n = 1 \ldots \ell\} \text{ TRAINING SET} \\ \text{DAL MODELLO} & \end{cases}$

$\begin{cases} g(\underline{x}) \text{ funzione parametrica } g(\underline{x}; \underline{c}) \\ g(\underline{x}) \text{ conseguenza del modello} \end{cases}$

parametri da trovare $<$ LINEARE (AFFINE) / NON LINEARE

CRITERIO DI COSTO (o DI REWARD / RITORNO)

REGRESSIONE SCALARE $\quad \ell = 6$



AFFINE

$g(x, \underline{c}) = c_1 x + c_0$

APPRENDIMENTO DAI DATI

$\{(x[1], d[1]), (x[2], d[2]), \ldots, (x[\ell], d[\ell])\}$

COSTO : MINIMI QUADRATI

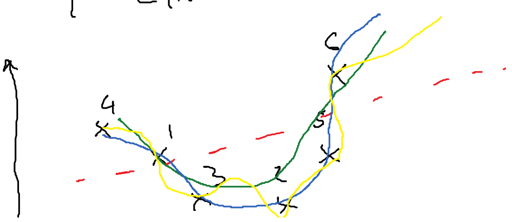
$$\underline{c}^o = \arg\min_{\underline{c}} \sum_{n=1}^{\ell} (y[n] - d[n])^2 = \arg\min_{(c_1, c_0)} \underbrace{\sum_{n=1}^{\ell} ((c_1 x[n] + c_0) - d[n])^2}_{\mathcal{E}(\underline{c})}$$

$$\frac{\partial \mathcal{E}(\underline{c})}{\partial c_0} = 2 \sum_{n=1}^{\ell} (c_1 x[n] + c_0 - d[n]) \cdot 1 = 0$$

$$\frac{\partial \mathcal{E}(\underline{c})}{\partial c_1} = 2 \sum_{n=1}^{\ell} (c_1 x[n] + c_0 - d[n]) x[n] = 0$$

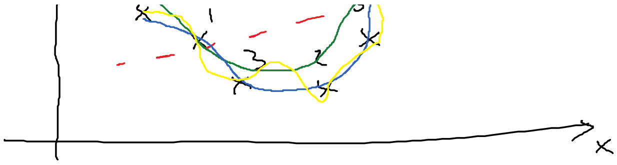
$$\begin{cases} c_1 \underbrace{\boxed{\sum_{n=1}^{\ell} x[n]}}_{\bar{x}} + c_0 = \underbrace{\boxed{\sum_{n=1}^{\ell} d[n]}}_{\bar{d}} \\ c_1 \underbrace{\boxed{\sum_{n=1}^{\ell} x[n]^2}}_{\overline{x^2}} + c_0 \underbrace{\boxed{\sum_{n=1}^{\ell} x[n]}}_{\bar{x}} = \underbrace{\boxed{\sum_{n=1}^{\ell} d[n] x[n]}}_{r_{dx}} \end{cases}$$

$$\begin{cases} c_1 \bar{x} + c_0 = \bar{d} \\ c_1 \overline{x^2} + c_0 \bar{x} = r_{dx} \end{cases} \Rightarrow (c_0^o, c_1^o)$$



$g(x, \underline{c}) = c_2 x^2 + c_1 x + c_0$ pol II ordine

$g(x, \underline{c}) = c_3 x^3 + c_2 x^2 + c_1 x + c_0$ pol III ordine



$$g(\underline{x},\underline{c}) = c_3 x^3 + c_2 x^2 + c_1 x + c_0 \quad \text{per ...}$$

$$= \underbrace{\left[c_3\ c_2\ c_1\ c_0\right]}_{C^T} \underbrace{\begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}}_{\underline{x}} = \underline{c}^T \underline{x} \qquad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$\underline{c}^o = \arg\min_{\underline{c}} \underbrace{\sum_{n=1}^{\ell} \left(\underline{c}^T \underline{x}[n] - d[n]\right)^2}_{\mathcal{E}(\underline{c})}$$

$$\frac{\partial \mathcal{E}(\underline{c})}{\partial c_i} = \cancel{2} \sum_{n=1}^{\ell} \left(\underline{c}^T \underline{x}[n] - d[n]\right) x_i[n] = 0 \qquad \forall\, i = 1,2,3,4$$

$$\sum_n \underline{c}^T \underline{x}[n]\, x_1[n] = \sum_n d[n]\, x_1[n]$$
$$\sum_n \underline{c}^T \underline{x}[n]\, x_2[n] = \sum_n d[n]\, x_2[n]$$
$$\sum_n \underline{c}^T \underline{x}[n]\, x_3[n] = \sum_n d[n]\, x_3[n]$$
$$\sum_n \underline{c}^T \underline{x}[n]\, x_4[n] = \sum_n d[n]\, x_4[n]$$

$$\begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \\ x_4[n] \end{bmatrix} \underline{x}_1$$

$$\boxed{\underbrace{\sum_{n=1}^{\ell} \underline{x}[n]\, \underline{x}^T[n]}_{4 \times 4} \underline{c} = \sum_n d[n]\, \underline{x}[n]}$$

$$\begin{bmatrix} 1 \\ x[n] \\ x^2[n] \\ x^3[n] \end{bmatrix} \left[1\ \ x[n]\ \ x^2[n]\right]$$

$$\left(\sum_n \begin{bmatrix} 1 \\ x[n] \\ x^2[n] \\ x^3[n] \end{bmatrix} \left[1\ x[n]\ x^2[n]\ x^3[n]\right]\right) \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \sum_n d[n] \\ \sum_n d[n]\, x[n] \\ \sum_n d[n]\, x^2[n] \\ \sum_n d[n]\, x^3[n] \end{bmatrix}$$

$$\begin{bmatrix} \sum_n 1 & \sum_n x[n] & \sum x^2[n] & \sum x^3[n] \\ \sum x[n] & \sum x^2[n] & \sum x^3[n] & \sum x^4[n] \\ \sum x^2[n] & \sum x^3[n] & \sum x^4[n] & \sum x^5[n] \\ \sum x^3[n] & \sum x^4[n] & \sum x^5[n] & \sum_n x^6[n] \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \downarrow \end{bmatrix} \rightarrow \underline{c}^o$$

REGRESSIONE LINEARE (CASO GENERALE)

$$N\Big\{\ \underline{x} = \begin{bmatrix} 1 \\ \underline{x}' \end{bmatrix} \rightarrow \boxed{\underline{\underline{C}}} \rightarrow \underline{y} \rightarrow \oplus \xleftarrow{+\, d[n]} \rightarrow \underline{\varepsilon}$$

$$\mathcal{E}(\underline{c}) = \sum_{n=1}^{\ell} \underline{\varepsilon}^T[n]\, \underline{\varepsilon}[n]$$

$$\underline{y} = \underline{\underline{C}}^T \underline{x} = \underline{g}(\underline{x};\underline{c})$$

$$M[\ ] = \underbrace{[\ M \cdots\ ]}^{N} [\ ]N$$

DAI DATI (DATA-DRIVEN)

TRAINING SET: $\{(\underline{x}[n]\ d[n]),\ n = 1,..,\ell\}$ ESEMPI $\Longrightarrow$ TEST SET: $\{(x[n]\ d[n])\ n = 1,..,T\}$

$$\underline{\underline{C}}^o = \arg\min \mathcal{E}(\underline{\underline{C}}) = \arg\min \sum_{n=1}^{\ell} \underline{\varepsilon}^T[n]\, \underline{\varepsilon}[n]$$

ESEMPI

$$\underline{\underline{C}}^{o} = \arg\min_{\underline{\underline{C}}} \mathcal{E}(\underline{\underline{C}}) = \arg\min_{\underline{\underline{C}}} \sum_{n=1}^{\ell} \underline{\varepsilon}^{T}[n]\,\underline{\varepsilon}[n]$$

$$= \arg\min_{\underline{\underline{C}}} \sum_{n=1}^{\ell} \left(\underline{d}[n] - \underline{\underline{C}}^{T}\underline{x}[n]\right)^{T}\left(\underline{d}[n] - \underline{\underline{C}}^{T}\underline{x}[n]\right)$$

$$\mathcal{E}(\underline{\underline{C}}) = \underbrace{\sum_n \underline{d}^{T}[n]\,\underline{d}[n]}_{\mathcal{E}_d} + \sum_n \underline{x}^{T}[n]\,\underline{\underline{C}}\,\underline{\underline{C}}^{T}\underline{x}[n] - \sum_n \underline{x}^{T}[n]\,\underline{\underline{C}}\,\underline{d}[n] - \sum_n \underline{d}^{T}[n]\,\underline{\underline{C}}^{T}\underline{x}[n]$$

$$\mathrm{tr}[AB] = \mathrm{tr}[BA]$$

$$= \mathcal{E}_d + \sum_n \mathrm{tr}\left[\underbrace{\underline{x}^{T}[n]\,C}_{A}\,\underbrace{C^{T}\underline{x}[n]}_{B}\right] - 2\sum_n \mathrm{tr}\left[\underline{x}^{T}[n]\,\underline{\underline{C}}\,\underline{d}[n]\right]$$

$$= \mathcal{E}_d + \sum_n \mathrm{tr}\left(\underline{\underline{C}}\,\underline{\underline{C}}^{T}\underline{x}[n]\,\underline{x}^{T}[n]\right) - 2\sum_n \mathrm{tr}\left(\underline{\underline{C}}\,\underline{d}[n]\,\underline{x}^{T}[n]\right)$$

$$= \mathcal{E}_d + \mathrm{tr}\Big(CC^{T}\underbrace{\sum_n x[n]\,x^{T}[n]}_{R_x}\Big) - 2\,\mathrm{tr}\Big(C\underbrace{\sum_n d[n]\,x^{T}[n]}_{R_{dx}}\Big)$$

$$\boxed{\mathcal{E}(\underline{\underline{C}}) = \mathcal{E}_d + \mathrm{tr}\left(CC^{T}R_x\right) - 2\,\mathrm{tr}\left(C R_{dx}\right)}$$

$X = \underline{\underline{C}}$ $\quad A = I$ $\quad B = R_x$

R_x

$\Uparrow$

$$R_x^{T} C + R_x C = 2\,\underline{\underline{R}}_x\,\underline{\underline{C}}$$

Table 4

$\phi(X)$	$\mathrm{d}\phi(X)$	$\mathrm{D}\phi(X)$	$\partial\phi(X)/\partial X$
$\mathrm{tr}\,AX$	$\mathrm{tr}\,A\mathrm{d}X$	$(\mathrm{vec}\,A')'$	A'
$\mathrm{tr}\,XAX'B$	$\mathrm{tr}(AX'B + A'X'B')\mathrm{d}X$	$(\mathrm{vec}(B'XA' + BXA))'$	$B'XA' + BXA$
$\mathrm{tr}\,XAXB$	$\mathrm{tr}(AXB + BXA)\mathrm{d}X$	$(\mathrm{vec}(B'X'A' + A'X'B'))'$	$B'X'A' + A'X'B'$

$$\mathrm{tr}\,(R_{dx}\,C)$$

A $\quad X$

grad.

R_{dx}^{T}

$$\nabla_{\underline{\underline{C}}}\,\mathcal{E}(\underline{\underline{C}}) = \cancel{2}\,\underline{\underline{R}}_x\,\underline{\underline{C}} - \cancel{2}\,\underline{\underline{R}}_{dx}^{T} = \underline{\underline{0}}$$

N $\quad N$ $\quad M$ $\quad N$ $\quad M$

$$\boxed{\underline{\underline{C}}^{o} = \underline{\underline{R}}_x^{-1}\,\underline{\underline{R}}_{dx}^{T}}$$

VALUTIAMO LA AUTOCORRELAZIONE DI $\underline{\varepsilon}$

$$\underline{\underline{R}}_{\varepsilon} = \sum_n \underline{\varepsilon}[n]\,\underline{\varepsilon}^{T}[n] = \sum_n \left(\underline{\underline{C}}^{T}\underline{x}[n] - \underline{d}[n]\right)\left(\underline{\underline{C}}^{T}\underline{x}[n] - \underline{d}[n]\right)^{T}$$

$$= \sum_n \underline{\underline{C}}^{T}\underline{x}[n]\,\underline{x}^{T}[n]\,\underline{\underline{C}} + \sum_n \underline{d}[n]\,\underline{d}^{T}[n] - \sum_n \underline{d}[n]\,\underline{x}^{T}[n]\,C$$

$$- \sum_n \underline{\underline{C}}^{T}\underline{x}[n]\,\underline{d}^{T}[n]$$

$$= \underline{\underline{C}}^{T}\underline{\underline{R}}_x\,\underline{\underline{C}} + \underline{\underline{R}}_d - R_{dx}C - C^{T}R_{dx}^{T}$$

IN CONDIZIONI DI OTTIMALITÀ $\quad C^{o} = R^{-1}R^{T}$

$$= \underline{\underline{C}}^T \underline{\underline{R}}_x \underline{\underline{C}} + \underline{\underline{R}}_d - R_{dx} C^o - C^T R_{dx}$$

IN CONDIZIONI DI OTTIMALITÀ $\quad \underline{\underline{C}}^o = \underline{\underline{R}}_x^{-1} R_{dx}^T$

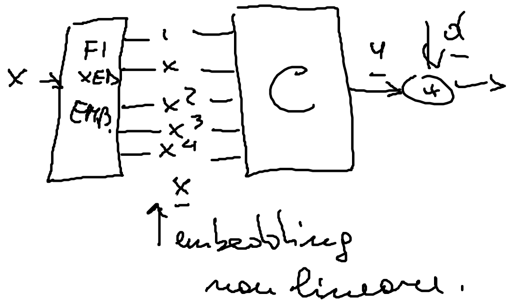
$$= R_{dx} \cancel{R_x^{-1}} \cancel{R_x} R_x^{-1} R_{dx}^T + \underline{\underline{R}}_d - R_{dx} R_x^{-1} R_{dx}^T - R_{dx} R_x^{-1} R_{dx}^T$$

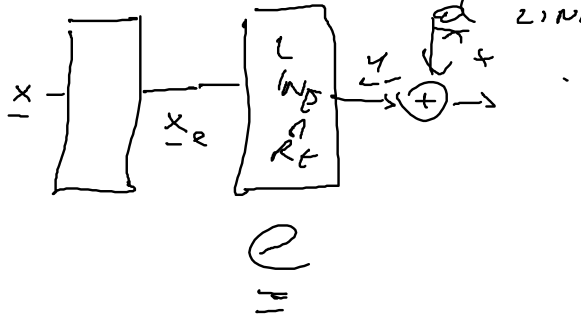
$$\underline{\underline{R}}_\varepsilon^o = \underline{\underline{R}}_d - R_{dx} R_x^{-1} R_{dx}^T$$

$$\mathcal{E}^o = tr(\underline{\underline{R}}_\varepsilon^o) = tr(\underline{R}_d) - tr(\underbrace{R_{dx} R_x^{-1}}_{C^T} R_{dx}^T)$$

$$= \mathcal{E}_d - \underline{\underline{C}}^{oT} \underline{\underline{R}}_{dx}$$

REGRESSIONE SU UN EMBEDDING NON LINEARE





The embedding couls be a whole network or a pre-processing block

Next case: We have available the moments and write the solution

$$\underline{x} = \begin{bmatrix} 1 \\ \underline{x}' \end{bmatrix} \to \boxed{\underline{C}} \to \underline{y} \to \oplus \to \underline{\varepsilon} \qquad (\underline{d} \text{ in ingresso al } \oplus)$$

$$\boxed{\underline{\underline{C}}^o = \underline{\underline{R}}_x^{-1} \underline{\underline{R}}_{dx}^T}$$

$$\boxed{\begin{aligned} \underline{\underline{R}}^o_\varepsilon &= \underline{\underline{R}}_d - R_{dx} R_x^{-1} R_{dx}^T \\ \mathcal{E}_o &= \mathcal{E}_d - \underline{\underline{C}}^{oT} \underline{\underline{R}}_{dx} \end{aligned}}$$

↘ dai dati

$$R_x = \sum_n \underline{x}[n]\,\underline{x}^T[n]$$

$$R_{dx} = \sum_n \underline{d}[n]\,\underline{x}^T[n]$$

$$\mathcal{E}_d = \sum_n \underline{d}^T[n]\,\underline{d}[n]$$

MINIMO ERRORE QUADRATICO

MSSE ← min, sum, square, error

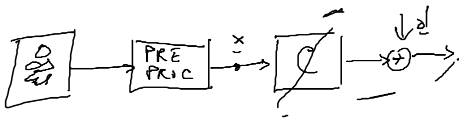
$$\begin{cases} R_x = E[\underline{x}\,\underline{x}^T] \\ R_{dx} = E[\underline{d}\,\underline{x}^T] \\ \mathcal{E}_d = E[\underline{d}^T\underline{d}] \end{cases}$$

⟹ stesse soluzioni formali

MMSE ← min, mean, square, error

$$\underline{\underline{C}}^o = \arg\min_{\underline{C}} E[\underline{\varepsilon}^T\underline{\varepsilon}]$$

MINIMO ERRORE QUADRATICO MEDIO



$$\underline{x} \to \boxed{g(\underline{x},\underline{c})} \to \underline{y} \to \oplus \to \underline{\varepsilon}, \qquad \underline{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_M \end{bmatrix}$$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} \qquad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix}$$

$E[\underline{d}\,|\,\underline{x}]$

$P_{\underline{X}\underline{D}}(\underline{x},\underline{d})$

$P_{\underline{D}|\underline{x}}(\underline{d}\,|\,\underline{x})$

$$\mathcal{E} = E\left[(\underline{D}-\underline{Y})^T(\underline{D}-\underline{Y})\right] = E[\underline{\varepsilon}^T\underline{\varepsilon}]$$

$$= \int_{\mathcal{X}\times\mathcal{D}} (\underline{d}-\underline{y})^T(\underline{d}-\underline{y})\, P_{\underline{X}\underline{D}}(\underline{x},\underline{d})\, d\underline{x}\, d\underline{d}$$

$$P_{\underline{X}\underline{D}}(\underline{x},\underline{d}) \to P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, P_{\underline{x}}(\underline{x})$$

$$= \int_{\mathcal{X}} P_{\underline{x}}(\underline{x}) \underbrace{\int_{\mathcal{D}} (\underline{d}-\underline{y})^T(\underline{d}-\underline{y})\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d}}_{\mathcal{E}(\underline{x})}\, d\underline{x}$$

$$\mathcal{E}(\underline{x}) = \int_{\mathcal{D}} (\underline{d}-\underline{y})^T(\underline{d}-\underline{y})\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d} = \int \underline{d}^T\underline{d}\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d} - 2\, \underline{y}^T \int_{\mathcal{D}} \underline{d}\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d} + \underline{y}^T\underline{y} \overbrace{\cancel{\int_{\mathcal{D}} P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d}}}^{1}$$

↑ minimizzare $\mathcal{E}(\underline{x})$

Riduciamo $\mathcal{E}(\underline{x})$ a forma canonica

$$\mathcal{E}(\underline{x}) = \left(\underline{y}^T - \int \underline{d}^T P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d}\right)\left(\underline{y} - \int \underline{d}\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d}\right) + \underbrace{\int \underline{d}^T\underline{d}\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d}} - \underbrace{\left\| \int \underline{d}\, P_{\underline{D}|\underline{x}}(\underline{d}|\underline{x})\, d\underline{d} \right\|^2}$$

$$\Rightarrow \mathcal{E}_{opt}(\underline{x}) =$$

$$\Rightarrow \mathcal{E}_{opt}(x) =$$

$$\boxed{\underline{y}^o = \int \underline{d}\, p_{\underline{D}|\underline{X}}(\underline{d}|\underline{x})\, d\underline{d} = E[\underline{d}\,|\,\underline{X}=\underline{x}]}$$

$$\boxed{g^{opt}(\underline{x}) = E[\underline{d}\,|\,\underline{X}=\underline{x}]}$$

$$\nabla_y y^T y = 2\underline{y}$$

PROVA ALTERNATIVA

$$\nabla_{\underline{y}} \mathcal{E}(\underline{x}) = \underline{0}$$

$$\mathcal{E}(\underline{x}) = \int \underline{d}^T \underline{d}\, p_{\underline{D}|\underline{X}}(\underline{d}|\underline{x})\, d\underline{d} - 2\underline{y}^T \int \underline{d}\, p_{D|X}(\underline{d}|\underline{x})\, d\underline{d} + \underline{y}^T \underline{y}$$

$$\Downarrow$$

$$-2\int \underline{d}\, p_{D|X}(\underline{d}|\underline{x})\, d\underline{d} + 2\underline{y} = \underline{0} \Rightarrow \boxed{\underline{y}^o = \int \underline{d}\, p_{D|X}(\underline{d}|\underline{x})\, d\underline{d}}$$

$$H = \left(\nabla_{\underline{y}} \nabla_{\underline{y}}^T\right) \mathcal{E}(\underline{x}) = 2\underline{I} > 0$$

(HESSIANO)

CASO GAUSSIANO

$\underline{x}$ $\quad \underline{d}$

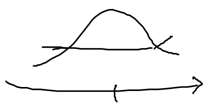
$E[\underline{d}\,|\,\underline{x}]$

$\bar{\underline{x}} = E[\underline{x}] \quad \bar{\underline{d}} = E[\underline{d}]$

$\underline{\underline{\Sigma}}_x = cov[\underline{x}] \quad \underline{\underline{\Sigma}}_d = cov[\underline{d}]$

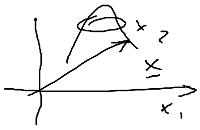
$$E[(\underline{X}-\bar{\underline{x}})(\underline{D}-\bar{\underline{d}})^T] = \underline{\underline{\Sigma}}_{\underline{X}\underline{D}}$$

$$\mathcal{N}(x;\bar{x},\sigma_x^2) = p_X(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}}\, e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}}$$



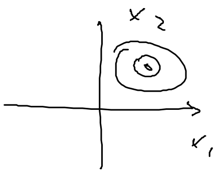
GAUSSIANA N-DIMENSIONALE

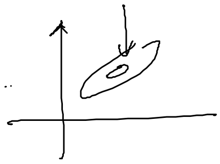
$\underline{x}$ vettore



$$p_{\underline{x}}(\underline{x}) = \mathcal{N}(\underline{x};\bar{\underline{x}},\underline{\underline{\Sigma}}_x) = \frac{1}{(2\pi)^{N/2}|\underline{\underline{\Sigma}}_x|}\, e^{-\frac{1}{2}(\underline{x}-\bar{\underline{x}})^T \underline{\underline{\Sigma}}_x^{-1}(\underline{x}-\bar{\underline{x}})}$$

determinante





$$\underline{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \qquad \underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} \qquad \underline{\underline{\Sigma}}_X \quad \underline{\underline{\Sigma}}_Y$$

$$\underline{\underline{\Sigma}}_{XY}$$

$$\underline{Z} = \begin{bmatrix} \underline{X} \\ \underline{Y} \end{bmatrix} \begin{matrix} \}N \\ \}M \end{matrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_N \\ \vdots \end{bmatrix} \Big\} N+M$$

$$\underline{Z} = \begin{bmatrix} \underline{X} \\ \underline{Y} \end{bmatrix} \begin{matrix} \}N \\ \}M \end{matrix} = \begin{bmatrix} \vdots \\ X_N \\ Y_1 \\ \vdots \\ Y_M \end{bmatrix} \Big\} N+M$$

$$\mathrm{Cov}[\underline{Z}] = E\left[(\underline{Z} - \bar{\underline{z}})(\underline{Z} - \bar{\underline{z}})^T\right] = E\left[\begin{bmatrix} \underline{X} - \bar{\underline{x}} \\ \underline{Y} - \bar{\underline{y}} \end{bmatrix} \begin{bmatrix} \underline{X}^T - \bar{\underline{x}}^T & \underline{Y}^T - \bar{\underline{y}}^T \end{bmatrix}\right]$$

$$= \begin{bmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_Y \end{bmatrix} \qquad L = N+M$$

$$p_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) = \frac{p_{\underline{X}\underline{Y}}(\underline{x},\underline{y})}{p_{\underline{Y}}(\underline{y})} = \frac{p_{\underline{Z}}(\underline{z})}{p_{\underline{Y}}(\underline{y})}$$

$$= \cancel{\frac{1}{(2\pi)^{L/2}|\Sigma_Z|} e^{-(x^T - \bar{x}^T ; y^T - \bar{y}^T)}} \quad \ldots\ldots =$$

$$p_{\underline{X}|\underline{Y}}(x|y) = \frac{1}{(2\pi)^{N/2}\left|\Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}\right|} e^{-\frac{1}{2}(\underline{x} - \bar{\bar{\underline{x}}})^T(\Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX})^{-1}(\underline{x} - \bar{\bar{\underline{x}}})}$$

DAGLI APPUNTI

$$\bar{\bar{x}} = \bar{x} + \Sigma_{XY}\Sigma_Y^{-1}(\underline{y} - \bar{\underline{y}})$$

$$\boxed{E[\underline{X}|\underline{Y} = \underline{y}] = \bar{x} + \Sigma_{XY}\Sigma_Y^{-1}(y - \bar{y})}$$

CASO GAUSSIANO $\qquad \underline{x} \qquad \underline{d} \qquad$ MMSE

$$\bar{\underline{x}} = E[\underline{x}] \qquad \bar{\underline{d}} = E[\underline{d}]$$

$$\Sigma_x = \mathrm{cov}[\underline{x}] \qquad \Sigma_d = \mathrm{cov}[\underline{d}]$$

$$E[(\underline{X} - \bar{\underline{x}})(\underline{D} - \bar{\underline{d}})^T] = \Sigma_{\underline{X}\underline{D}}$$

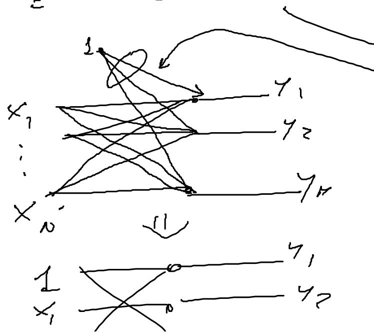
$$g^\circ(\underline{x}) = E[\underline{d}|\underline{x}]$$

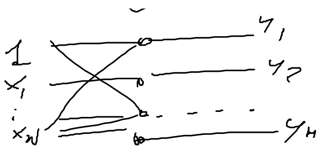
$$\downarrow$$

$$\underline{g}^\circ(\underline{x}) = E[\underline{d}|\underline{x}] = \bar{\underline{d}} + \Sigma_{DX}\Sigma_X^{-1}(\underline{x} - \bar{\underline{x}})$$

AFFINE!
(LINEARE SE $\bar{\underline{d}} = \underline{0}$)

$$\rightarrow \underbrace{\bar{\underline{d}} - \Sigma_{DX}\Sigma_X^{-1}\bar{\underline{x}}} + \underbrace{\Sigma_{DX}\Sigma_X^{-1}\underline{x}}$$

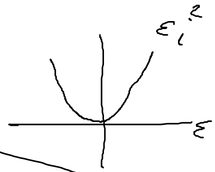




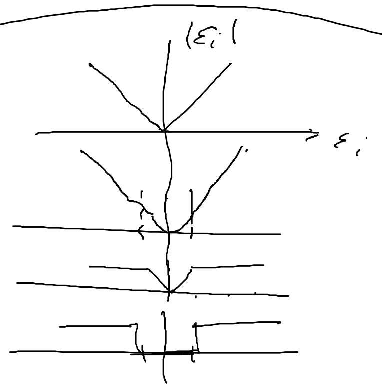
$$\mathcal{E} = E[\varepsilon^T \varepsilon] \; E[(\underline{x}-\underline{d})^T(\underline{x}-\underline{d})] \quad MMSE$$

$$= \sum_n (x[n]-d[n])^T (x[n]-d[n]) \quad MSSE$$

$n \in$ Training set.



Altre funzioni di costo



$P_{\underline{x}\,\underline{D}}(\underline{x},\underline{d}) \qquad P_{\underline{D}/\underline{x}}(\underline{d}|\underline{x})\, p_{\underline{x}}(\underline{x})$

MAP MAXIMUM APOSTERIORE ESTIMATE

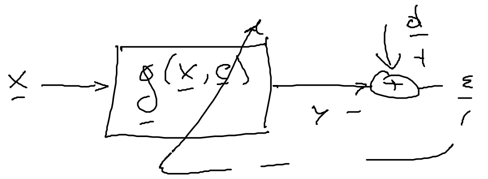
$\underline{y}^o = \arg\max_{\underline{d}} P_{\underline{D}/\underline{x}}(\underline{d}/\underline{x})$ $\Rightarrow$ Se $\underline{X}$ e $\underline{D}$ sono congiuntamente gaussiane.

$\Rightarrow \underline{y}^o = E[\underline{d}|\underline{x}]$ = funzione lineare

APPROCCIO ADATTATIVO

DATA DRIVEN

$\{(x[n]\, d[n]),\ n=1\ldots L\}$



PER CAPIRE : "Semplice" Problema $g(\underline{x},\underline{c}) = \underline{c}^T\underline{x}$

$E[\underline{x}]=0$ Noti $\underline{\underline{R}}_x$ $r_{\underline{x}d} = E[\underline{x}d]$ d scalare.

$\underline{c}^o = \underline{\underline{R}}_x^{-1} r_{\underline{x}d}$

MMSE

$$\mathcal{E}(\underline{c}) = E[(\underline{c}^T\underline{x} - d)^2]$$

$$= E[c^T x x^T c - 2\underline{c}^T\underline{x}d + d^2] = c^T R_x c - 2c^T \underline{r}_{xd} + \mathcal{E}_d$$

$\nabla_{\underline{x}}\, \underline{x}^T \underline{\underline{A}}\, \underline{x} = 2\underline{\underline{A}}\,\underline{x}$

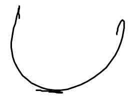
$\nabla_{\underline{x}}\, \underline{c}^T\underline{x} = \underline{c}$

$C[k] \longrightarrow$

Muoversi in direzione opposta al gradiente

$$\underline{c}[k+1] = \underline{c}[k] - \mu \nabla_{\underline{c}} \mathcal{E}(\underline{c}[k])$$

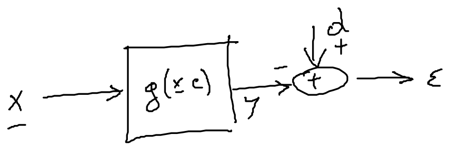
$$\nabla_{\underline{c}} \mathcal{E}(\underline{c}) = 2\underline{R}_x \underline{c} - 2\underline{r}_{xd}$$



punto di equilibrio

$$\underline{c}[k+1] = \underline{c}[k] - \mu \underbrace{\left(\underline{R}_x \underline{c}[k] - \underline{r}_{xd}\right)}_{=0} \Longrightarrow$$

AL G. AL GRADIENTE



$$\{(\underline{x}[n], d[n]),\ n = 1 \dots \ell\}$$

TRAINING SET
INSIEME DI APPRENDIMENTO

$$\varepsilon[n] = d[n] - g(\underline{x}[n], \underline{c})$$

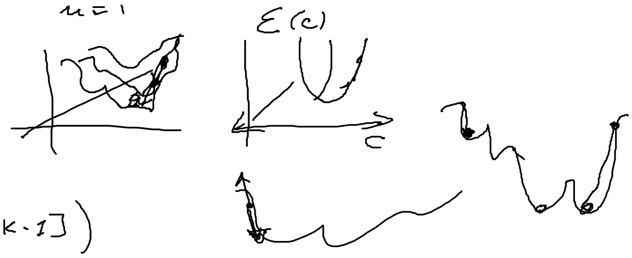
Minimizzare

$$\mathcal{E}(\underline{c}) = \sum_{n=1}^{\ell} \left(d[n] - g(\underline{x}[n], \underline{c})\right)^2$$

$$\underline{c}^o = \arg\min_{\underline{c}} \mathcal{E}(\underline{c}) = \arg\min_{\underline{c}} \sum_{n=1}^{\ell} \left(d[n] - g(\underline{x}[n], \underline{c})\right)^2$$

STRATEGIA AL GRADIENTE

$\underline{c}[0]$ cond. iniziale



$$\underline{c}[k] = \underline{c}[k-1] - \mu \nabla_{\underline{c}} \mathcal{E}(\underline{c}[k-1])$$

$$\begin{bmatrix} \frac{\partial \mathcal{E}}{\partial c_1} \\ \vdots \\ \frac{\partial \mathcal{E}}{\partial c_N} \end{bmatrix}^T = \nabla_{\underline{c}} \mathcal{E}(\underline{c}) = 2 \sum_{n=1}^{\ell} \left(d[n] - g(\underline{x}[n], \underline{c})\right)\left(-\nabla_{\underline{c}} g(\underline{x}[n], \underline{c})\right)$$

$$\underline{c}[k] = \underline{c}[k-1] - \mu \underbrace{\sum_{n=1}^{\ell} \left(d[n] - g(\underline{x}[n], \underline{c}[k-1])\right)\left(-\nabla_{\underline{c}} g(\underline{x}[n], \underline{c}[k-1])\right)}_{\text{su tutto il training set.}}$$

BATCH

CASO LINEARE

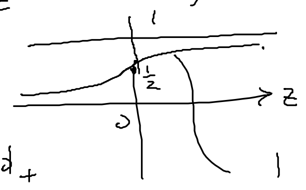
$$g(\underline{x}, \underline{c}) = \underline{c}^T \underline{x} \qquad \nabla_c g(x, c) = \underline{x}$$

$$\underline{c}[k] = \underline{c}[k-1] + \mu \sum_{n=1}^{\ell} \left(d[n] - \underbrace{\underline{c}^T[k-1]\underline{x}[n]}_{\underline{x}^T[n]\,\underline{c}[k-1]}\right) \underline{x}[n]$$

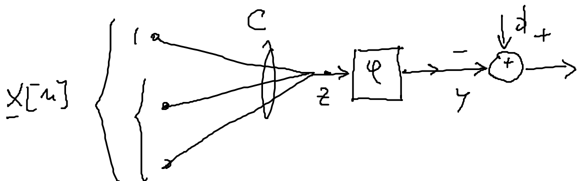
$$= \underline{c}[k-1] + \mu \sum_{n=1}^{\ell} \left(d[n]\underline{x}[n] - \underline{x}[n]\underline{x}[n]^T \underline{c}[k-1]\right)$$

$$= \underline{c}[k-1] + \mu \left(\underline{r}_{dx} - \underline{\underline{R}}_x \underline{c}[k-1]\right)$$

UN PRIMO CASO NONLINEARE



$$0 < d[n] < 1$$



$$\varphi(z) = \frac{1}{1 + e^{-z}}$$

FUNZIONE LOGISTICA

$$y = \varphi(\underline{c}^T \underline{x})$$

$$\{(\underline{x}[n], d[n]),\ n = 1 \dots \ell\}$$

$$\underbrace{y = \varphi(\underline{c}^T \underline{x})}_{g(\underline{x},\,\underline{c})} \qquad \{(\underline{x}[n], d[n]),\ n = 1 \dots \ell\}$$

$$\nabla_{\underline{c}}\, \varphi(\underline{c}^T \underline{x}) = \varphi'(\underline{c}^T \underline{x})\, \underline{x} \qquad \varphi'(z) = \frac{+e^{-z}}{(1+e^{-z})^2} \qquad \boxed{\varphi'(z) = \varphi(z)(1-\varphi(z))}$$

$$= \varphi(c^T x)\left(1 - \varphi(c^T x)\right)\underline{x} = \underbrace{z(1-z)\,\underline{x}}$$

PROVA:

$$\varphi(z)(1-\varphi(z)) = \frac{1}{1+e^{-z}}\left(1 - \frac{1}{1+e^{-z}}\right) = \frac{\cancel{1}+e^{-z}\cancel{-1}}{(1+e^{-z})^2}$$

$$\frac{1}{1+e^{-z}}\,\frac{e^{-z}+1-1}{(1+e^{-z})} = \frac{1}{1+e^{-z}}\left(1-\frac{1}{1+e^{-z}}\right)$$

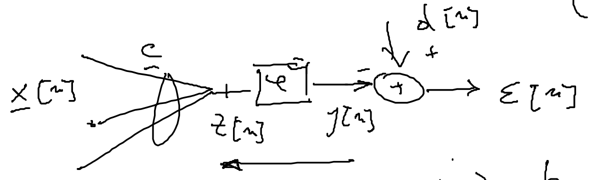
$$\underline{C}[k] = \underline{C}[k-1] - \mu \nabla_{\underline{c}}\, \mathcal{E}(\underline{c})$$

$$C[k] = C[k-1] - \mu \nabla_C \sum_{n=1}^{\ell}\left(d[n] - \varphi\left(\underline{c}^T[k-1]\, \underline{x}[n]\right)\right)^2$$

$$C[k] = \underline{c}[k-1] + \mu \underbrace{\sum_{n=1}^{\ell}\left(d[n] - \varphi\left(c^T[k-1]\,\underline{x}[n]\right)\right)\varphi\left(c^T[k-1]x[n]\right)\left(1-\varphi\left(c^T[k-1]x[n]\right)\right)\underline{x}[n]}_{\text{BATCH}}$$

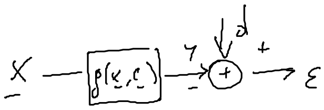
$$C[k-1] + \mu \sum_{n=1}^{\ell}\left(d[n] - \varphi(z[n])\right)\underbrace{\varphi(z[n])}_{y[n]}\left(1-\varphi(z[n])\right)\underline{x}[n] \qquad z[n]$$

$$c[k-1] + \mu \underbrace{\sum_{n=1}^{\ell}\underbrace{\left(d[n] - y[n]\right)}_{\varepsilon[n]}\underbrace{y[n](1-y[n])}_{\varphi'}\,\underline{x}[n]}$$



backpropagation dell'errore

$$\underline{C}[k] = \underline{C}[k-1] - \mu\, \nabla_{\underline{c}}\, \mathcal{E}(\underline{c}[k-1])$$



$$\underline{C}[k] = \underline{c}[k-1] - \mu \underbrace{\sum_{n=1}^{\ell}\left(d[n] - g(\underline{x}[n], \underline{c}[k-1])\right)\left(-\nabla_{\underline{c}}\, g(\underline{x}[n], \underline{c}[k-1])\right)}_{\text{su tutto il training set}}$$

BATCH → Si aggiorna una sola volta alla fine di ogni epoca (sulla media)

APPROCCIO STOCASTICO

$$\underline{C}[k] = \underline{C}[k-1] - \mu\, \nabla_{\underline{c}}\, \mathcal{E}(\underline{c}[k-1])$$

$$\underline{C}[k] = \underline{c}[k-1] - \mu \underbrace{\cancel{\sum_{n=1}^{\ell}}\left(d[n] - g(\underline{x}[n], \underline{c}[k-1])\right)\left(-\nabla_{\underline{c}}\, g(\underline{x}[n], \underline{c}[k-1])\right)}_{\text{su tutto il training set}}$$

EPOCH
Una presentazione completa del training set

$$\underline{C}[n] = \underline{c}[n-1] + \mu\left(d[n] - g(x[n], c[n-1])\right)\nabla_{\underline{c}}\, g(x[n], c[n-1])$$

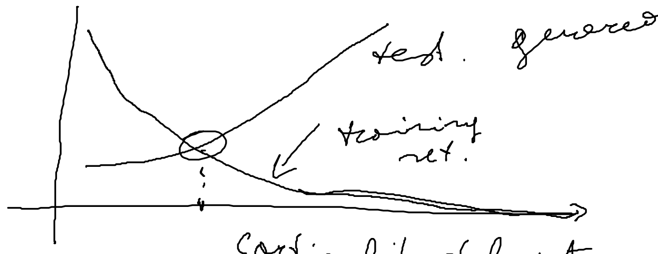
si aggiorna ad ogni esempio !!

PROBLEMA

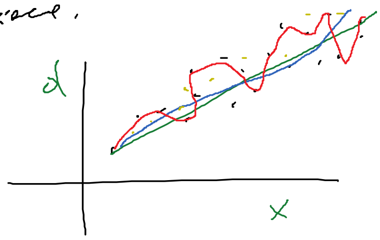
TRAINING SET | TEST SET | VALIDATION SET

$\{(x[n], d[n]), n = 1 \dots \ell\}$ | $\{(x[n], d[n]) \; n = 1, \dots \ell_T\}$ | $\{(x[n], d[n]), n = 1, \dots \ell_V\}$

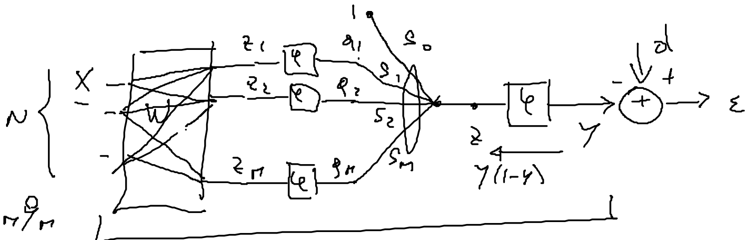
diversi.



Cardinalità del sistema
(gradi di libertà)



Esempio



$z = s_0 + s_1 q_1 + s_2 q_2 + \dots + s_M q_M$

$c = \{\underline{W}, \underline{s}\}$

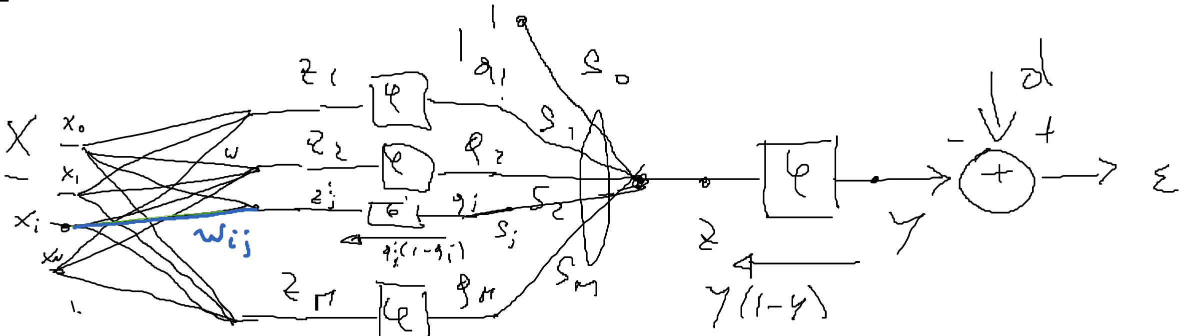
$y(\underline{x}; \underline{c})$

$\varepsilon = d - y = d - \varphi(z)$

$$\nabla_{\underline{c}} \varepsilon \quad \begin{bmatrix} \frac{\partial \varepsilon}{\partial s_0} \\ \frac{\partial \varepsilon}{\partial s_1} \\ \vdots \\ \frac{\partial \varepsilon}{\partial s_M} \end{bmatrix} = \begin{bmatrix} -\frac{\partial y}{\partial s_0} \\ -\frac{\partial y}{\partial s_1} \\ \vdots \\ -\frac{\partial y}{\partial s_M} \end{bmatrix} = \begin{bmatrix} -y(1-y)\frac{\partial z}{\partial s_0} \\ -y(1-y)\frac{\partial z}{\partial s_1} \\ \vdots \\ -y(1-y)\frac{\partial z}{\partial s_M} \end{bmatrix} = \begin{bmatrix} -y(1-y) \\ -y(1-y)q_1 \\ -y(1-y)q_2 \\ \vdots \\ -y(1-y)q_M \end{bmatrix}$$

$\underline{s}[k] = \underline{s}[k-1] + \mu \; \varepsilon[k] y[k](1 - y[k]) \underline{q}[k]$ ← stocastico

$$\underline{S}[k] = \underline{S}[k-1] + \mu \sum_{n=1}^{\ell} \left(d[n] - g(\underline{x}[n], c[k-1])\right) y[n](1-y[n]) \underline{q}[n]$$



$$\frac{\partial \varepsilon}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}}(d - y) = -\frac{\partial y}{\partial w_{ij}} = -\frac{\partial \varphi(z)}{\partial w_{ij}} = -\varphi'(z)\frac{\partial z}{\partial w_{ij}}$$

$$= -\varphi'(z)\frac{\partial}{\partial w_{ij}}(s_0 + s_1 q_1 + s_j q_j + s_M q_M) =$$

$$= -\varphi(z) \frac{\partial}{\partial w_{ij}} \left(S_0 + S_1 q_1 + s_j q_j + \dots + s_M q_M \right) =$$

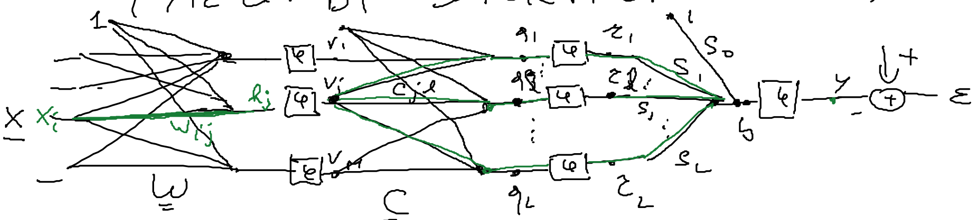
$$= -\varphi'(z)\, s_j \frac{\partial q_j}{\partial w_{ij}} = -\varphi'(z)\, s_j \frac{\partial \varphi(z_j)}{\partial w_{ij}} = -\varphi'(z)\, s_j\, q_j (1-q_j) \frac{\partial z_j}{\partial w_{ij}}$$

$$= -\varphi'(z)\, s_j\, q_j (1-q_j)\, x_i$$

$$= \underline{-\varphi(1-\varphi)\, s_j\, q_j (1-q_j)}\, \underline{x_i}$$

CALCOLO DEL GRADIENTE (ALG. DI BACKPROPAGATION)

$$\varphi'(x) = \varphi(x)(1-\varphi(x))$$



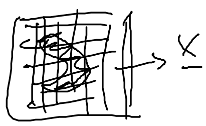
$$\frac{\partial y}{\partial w_{ij}} = \frac{\partial y}{\partial b}\frac{\partial b}{\partial w_{ij}} = y(1-y)\frac{\partial b}{\partial w_{ij}} = y(1-y)\sum_{\ell=1}^{L}\underbrace{\frac{\partial b}{\partial z_\ell}}_{s_\ell}\frac{\partial z_\ell}{\partial w_{ij}}$$

$$= y(1-y)\sum_{\ell=1}^{L} s_\ell \frac{\partial z_\ell}{\partial q_\ell}\frac{\partial q_\ell}{\partial w_{ij}} = y(1-y)\sum_{\ell=1}^{L} s_\ell z_\ell(1-z_\ell)\frac{\partial q_\ell}{\partial w_{ij}}$$

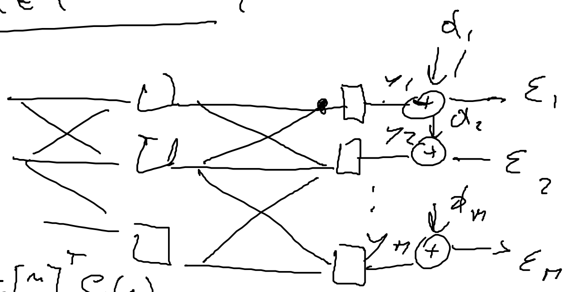
$$= y(1-y)\sum_{\ell=1}^{L} s_\ell z_\ell(1-z_\ell)\frac{\partial q_\ell}{\partial v_j}\frac{\partial v_j}{\partial w_{ij}} = y(1-y)\sum_{\ell=1}^{L} s_\ell z_\ell(1-z_\ell)\, c_{j\ell}\frac{\partial v_j}{\partial h_j}\frac{\partial h_j}{\partial w_{ij}}$$

$$= y(1-y)\sum_{\ell=1}^{L} s_\ell z_\ell(1-z_\ell)\, c_{j\ell}\, v_j(1-v_j)\, x_i$$

MNIST



$d \in \{0\,1\,2 \dots 9\} \longrightarrow$ one-hot



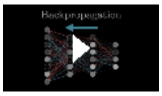
d			
1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

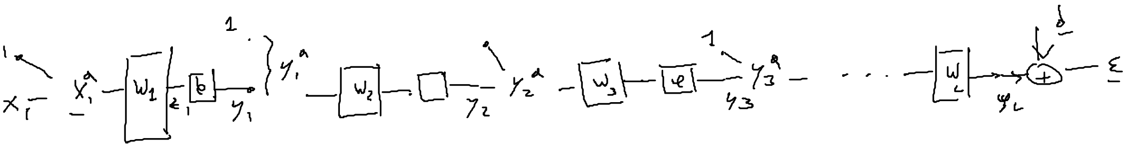
$$\mathcal{E}(\underline{c}) = \sum_{n=1}^{\ell} \underline{\varepsilon}[n]^T \underline{\varepsilon}(n)$$

$$[\varepsilon_1[n]\ \varepsilon_2[n] \dots \varepsilon_M[n]]\begin{bmatrix}\varepsilon_1[n]\\ \varepsilon_2[n]\\ \vdots\\ \varepsilon_M[n]\end{bmatrix} = \sum_{n=1}^{\ell}\sum_{i=1}^{M}\varepsilon_i[n]$$

Bella animazione grafica del backpropagation

[What is backpropagation really doing? | Chapter 3, Deep learning]

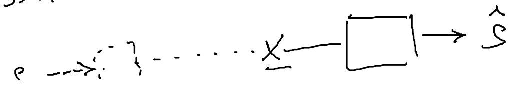




N $\quad N_1 \quad N_2 \quad \dots$

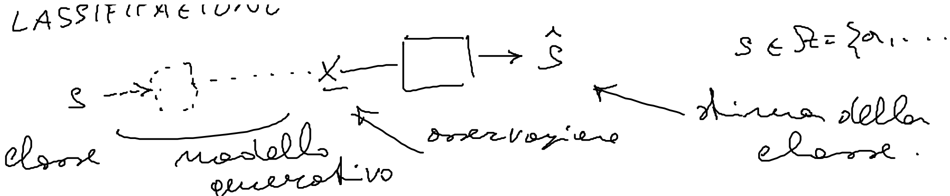
$$\underline{X} \longrightarrow X_1 \longrightarrow X_1^a \xrightarrow{W_1} z_1 \xrightarrow{\varphi} y_1 \xrightarrow{a} y_1^a \longrightarrow$$

CLASSIFICAZIONE



$s \in \Omega = \{a_1 \dots a_M\}$

CLASSIFICAZIONE



$s \in \Omega = \{a_1 \dots a_M\}$

$\underline{x} \in \mathcal{X} \subseteq \mathbb{R}^N$

Prob. a priori

$$f_{\underline{X}S}(\underline{x}, s) = f_{\underline{X}}(\underline{x}|s)\, \pi_S(s)$$

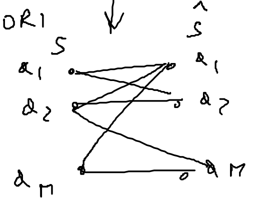
↑ informazione a priori nella classe

↑ funzione di verosimiglianza

OBIETTIVO:

Minimizzare $P(e) = P_r\{\hat{s} \neq s\}$

MATRICE DI CONFUSIONE



TESI: $\text{Min}\, P(e) \longleftrightarrow \text{Max}\ P_r\{s|\underline{x}\}$ PROBABILITA' A POSTERIORI

CRITERIO $\boxed{\text{MAP}}$ o CRITERIO A MASSIMA PROB. A POSTERIORI

$$P(e) = 1 - P(c) = 1 - \int_{\mathcal{X}} \underline{P(c|\underline{x})}\, f_{\underline{x}}(\underline{x})\, d\underline{x}$$

$$= \int_{\mathcal{X}} p(e|x)\, f_x(x)\, dx$$

osservo $\underline{x}$ → $p(a_1|\underline{x})$, $p(a_2|\underline{x})$, ⋮, $p(a_M|\underline{x})$ ⟷ max

MAP

$$\hat{s} = \arg\max_s\, p(s|\underline{x}) = \arg\max_s \frac{f_{x|s}(x|s)\,p(s)}{f_x(x)}$$

$$= \arg\max_s\, f_{\underline{x}|s}(\underline{x}|s)\,p(s)$$

$\underline{p(s) \text{ Uniforme}}$; $\underline{\text{non ne ho conoscenza}}$

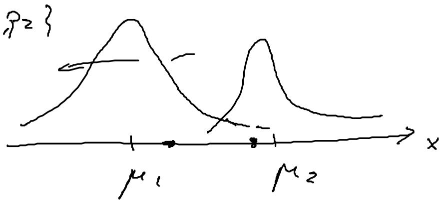
$$\Longrightarrow \arg\max_s\, f_{\underline{x}|s}(\underline{x}|s)$$

CRITERIO A MV MASSIMA VEROSIMIGLIANZA

ML - MAXIMUM LIKELIHOOD

Esempio $M = 2$ $s \in \Omega = \{a_1, a_2\}$ $\pi = \{p\ \ 1-p\}$
$\{p_1\ \ p_2\}$

$$f_{x|s}(x|s) = \begin{cases} N(x; \mu_1, \sigma_1^2) & s = a_1 \\ N(x; \mu_2, \sigma_2^2) & s = a_2 \end{cases}$$



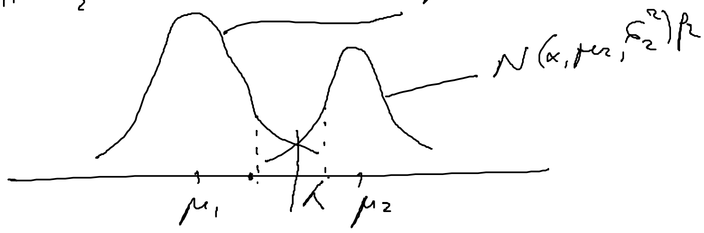
$$\begin{cases} \dfrac{1}{\sqrt{2\pi\sigma_1^2}}\, e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} & s = a_1 \\[2ex] \dfrac{1}{\sqrt{2\pi\,\sigma_2^2}}\, e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} & s = a_2 \end{cases}$$

$f(x|a_1)p \overset{a_1}{\gtrless} f(x|a_2)p$

$\sqrt{2\pi\,\sigma_2^2}$

$$f_x(x|a_1)\,p_1 \underset{a_2}{\overset{a_1}{\gtrless}} f_x(x|a_2)\,p_2$$

$$\frac{1}{\sqrt{\cancel{2\pi}\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} p_1 \underset{a_2}{\overset{a_1}{\gtrless}} \frac{1}{\sqrt{\cancel{2\pi}\,\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} p_2$$

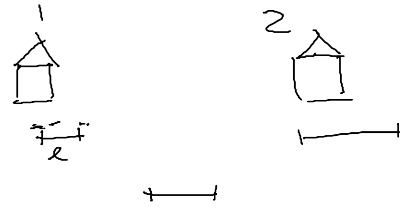


Lezione 20

lunedì 29 novembre 2021 14.06

$f_x(x|s)$

$s \dashrightarrow$ [MOD. GENERATIVO] $\dashrightarrow x \longrightarrow \boxed{CLASS} \longrightarrow \hat{s} \in \Omega = \{a_1, a_2\}$

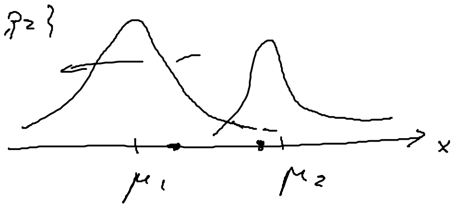


MODEL-BASED

Esempio $M=2 \quad s \in \Omega = \{a_1, a_2\} \quad \pi = \{p \;\; 1-p\}$

$\{p_1 \;\; p_2\}$

$$f_{x|s}(x|s) = \begin{cases} N(x;\mu_1, \sigma_1^2) & s = a_1 \\ N(x;\mu_2, \sigma_2^2) & s = a_2 \end{cases}$$

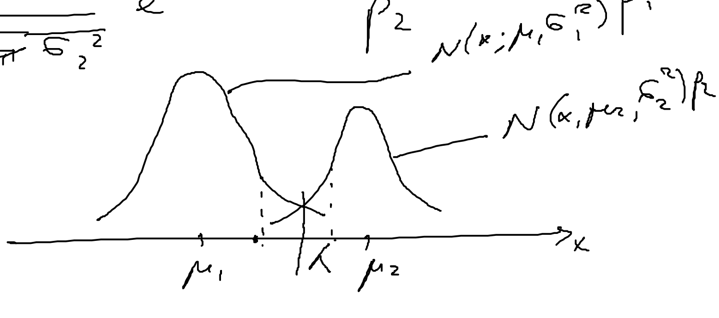


$$\begin{cases} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} & s = a_1 \\ \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} & s = a_2 \end{cases}$$

MAP

$$f_x(x|a_1)\, p_1 \underset{a_2}{\overset{a_1}{\gtrless}} f_x(x|a_2)\, p_2$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} p_1 \underset{a_2}{\overset{a_1}{\gtrless}} \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} p_2$$

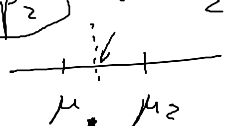


$x \longrightarrow \boxed{\text{RIV. DI SOGLIA}}_{\lambda} \longrightarrow \; < \nearrow a_1, \; > \searrow a_2$

$\downarrow \ln$

$$\frac{1}{2}\ln\frac{1}{\sigma_1^2} - \frac{(x-\mu_1)^2}{2\sigma_1^2} + \ln p_1 \underset{a_2}{\overset{a_1}{\gtrless}} \frac{1}{2}\ln\frac{1}{\sigma_2^2} - \frac{(x-\mu_2)^2}{2\sigma_2^2} + \ln p_2$$

$\lambda = \frac{\mu_1+\mu_2}{2}$



CASO PARTICOLARE (ML) $\begin{cases} p_1 = p_2 = \frac{1}{2} \\ \sigma_1^2 = \sigma_2^2 = \sigma \end{cases}$ $\Rightarrow -(x-\mu_1)^2 \underset{a_2}{\overset{a_1}{\gtrless}} -(x-\mu_2)^2$

$$(x-\mu_1)^2 \underset{a_1}{\overset{a_2}{\gtrless}} (x-\mu_2)^2 \Rightarrow \text{regola a minima distanza}$$

Caso particolare $\{\sigma_1^2 = \sigma_2^2 = \sigma^2$

caso particolare $\left\{ \sigma_1^2 = \sigma_2^2 = \sigma^2 \right.$

$$-\frac{(x-\mu_1)^2}{2\sigma^2} + \ln p_1 \underset{a_2}{\overset{a_1}{\gtrless}} -\frac{(x-\mu_2)^2}{2\sigma^2} + \ln p_2$$

$$-(x-\mu_1)^2 + 2\sigma^2 \ln p_1 \underset{a_2}{\overset{a_1}{\gtrless}} -(x-\mu_2)^2 + 2\sigma^2 \ln p_2$$

$$\cancel{-x^2} - \underline{\mu_1^2} + 2x\mu_1 + 2\sigma^2 \ln p_1 \underset{a_2}{\overset{a_1}{\gtrless}} \cancel{-x^2} - \underline{\mu_2^2} + 2x\mu_2 + 2\sigma^2 \ln p_2$$

$$2x(\mu_1-\mu_2) \underset{a_2}{\overset{a_1}{\gtrless}} \mu_1^2-\mu_2^2 + 2\sigma^2 \ln\frac{p_2}{p_1}$$

$$x \underset{a_2}{\overset{a_1}{\gtrless}} \frac{\mu_1^2-\mu_2^2}{2(\mu_1-\mu_2)} + \frac{\cancel{2}\sigma^2}{\cancel{2}(\mu_1-\mu_2)} \ln\frac{p_2}{p_1}$$

$$x \underset{a_2}{\overset{a_1}{\gtrless}} \overbrace{\frac{\mu_1+\mu_2}{2} + \frac{\sigma^2}{\mu_1-\mu_2} \ln\frac{p_2}{p_1}}^{\alpha}$$

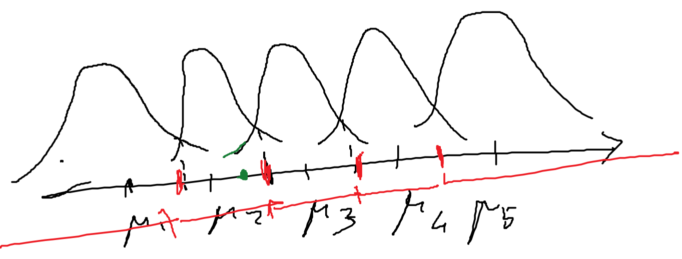
$M = 5 \qquad x \in \mathbb{R} \qquad p_i = \frac{1}{5} \quad i = 1..5 \qquad \Omega = \{a_1, a_2, a_3, a_4, a_5\}$

$f_x(x|a_1) = \mathcal{N}(x, \mu_1, \sigma^2)$

$f_x(x, a_2) = \mathcal{N}(x, \mu_2, \sigma^2)$

$\vdots$

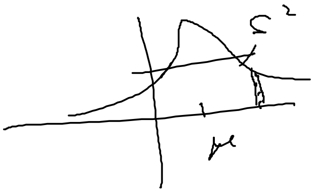
$f_x(x, a_5) = \mathcal{N}(x, \mu_5, \sigma^2)$



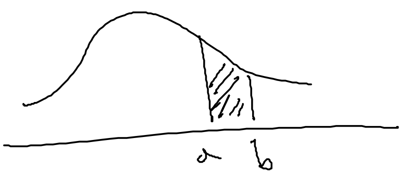
$\hat{s} = a_i$ se $f(x|a_i) > f(x|a_j) \quad j \neq i$

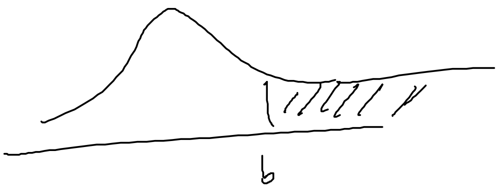
RICHIAMO

pdf gaussiana



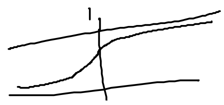
$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \mathcal{N}(x;\mu,\sigma^2)$$



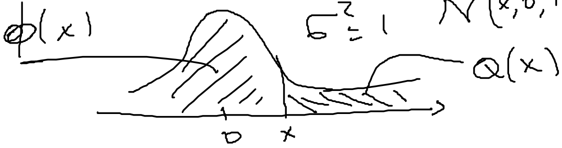


GAUSSIANA NORMALIZZATA

$$\dots = \frac{1}{\dots} e^{-\frac{x^2}{2}} \qquad x$$

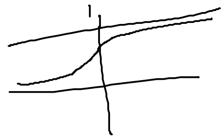


GAUSSIANA NORMALIZZATA



$$\sigma^2 = 1 \qquad N(x,0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\phi(x) = \int_{-\infty}^{x} N(\xi, 0, 1)\, d\xi$$

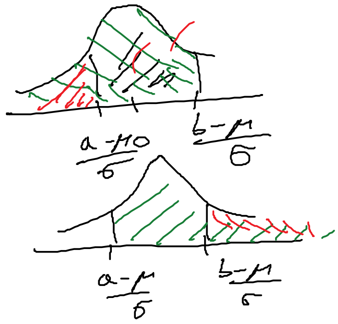


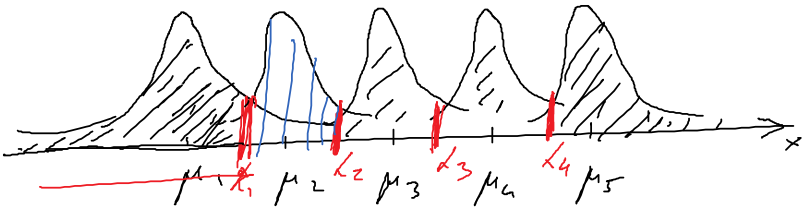
$$\int_a^b \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}\,\cancel{\sigma}} e^{-\frac{\xi^2}{2}} \cancel{\sigma}\, d\xi$$

$$\xi = \frac{x-\mu}{\sigma}$$

$$x = \sigma\xi + \mu \Rightarrow dx = \sigma\, d\xi$$

$$\phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right) = Q\left(\frac{a-\mu}{\sigma}\right) - Q\left(\frac{b-\mu}{\sigma}\right)$$





$$P(e) = 1 - P(c) = 1 - \sum_{i=1}^{5} P_z(c|a_i) p_i =$$

$$= 1 - P_z(c|a_1)p_1 - P_z(c|a_2)p_2 - P_z(c|a_3)p_3 - P_z(c|a_4)p_4 - P(c|a_5)p_5$$

$$= 1 - \left(1 - Q\left(\frac{\lambda_1 - \mu_1}{\sigma}\right)\right)p_1 - \left(1 - 2Q\left(\frac{\lambda_2-\mu_2}{\sigma}\right)\right)p_2 \quad \cdot - -$$

$$= 1 - \left(1 - Q\left(\frac{\lambda_1-\mu_1}{\sigma}\right)\right)(p_1 + p_5) - \left(1 - 2Q\left(\frac{\lambda_2-\mu_2}{\sigma}\right)\right)(p_2+p_3+p_4)$$

$$\underline{x} \in \mathbb{R}^N \qquad \Omega = \{a_1 \dots a_M\}$$

$$f_{\underline{x}}(\underline{x}|a_i) = N(\underline{x}; \underline{\mu}_i, \underline{\underline{\Sigma}}_i) = \frac{1}{(2\pi)^{N/2}|\underline{\underline{\Sigma}}_i|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T \underline{\underline{\Sigma}}_i^{-1}(\underline{x}-\underline{\mu}_i)}$$

MAP $\quad \hat{e} = a_i \quad f_{\underline{x}}(\underline{x}|a_i)p_i > f_{\underline{x}}(\underline{x}|a_j)p_j \quad j \neq i$

$$\frac{1}{\cancel{(2\pi)^{N/2}}|\underline{\underline{\Sigma}}_i|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T \underline{\underline{\Sigma}}_i^{-1}(\underline{x}-\underline{\mu}_i)} p_i > \frac{1}{\cancel{(2\pi)^{N/2}}|\underline{\underline{\Sigma}}_j|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_j)^T \underline{\underline{\Sigma}}_j^{-1}(\underline{x}-\underline{\mu}_j)} p_j$$

$$\frac{1}{\cancel{(2\pi)^{N/2}}|\underline{\underline{\Sigma}}_i|^{1/2}} e \quad p_i > \frac{1}{\cancel{(2\pi)^{N/2}}|\underline{\underline{\Sigma}}_j|^{1/2}} e \quad p_j$$

per capire, scegli fidelmente

$$MV \begin{cases} p_i = \frac{1}{M} & i=1..M \\ \underline{\underline{\Sigma}}_i = \underline{\underline{\Sigma}} & i=1..M \end{cases}$$

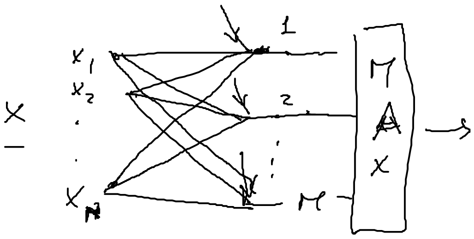
$$e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T\underline{\underline{\Sigma}}^{-1}(\underline{x}-\underline{\mu}_i)} > e^{-\frac{1}{2}(\underline{x}^T-\underline{\mu}_j^T)\underline{\underline{\Sigma}}^{-1}(x-\mu_j)}$$

$\Downarrow$ ln

$$+\cancel{\tfrac{1}{2}}(\underline{x}-\underline{\mu}_i)^T\Sigma^{-1}(\underline{x}-\underline{\mu}_i) < +\cancel{\tfrac{1}{2}}(\underline{x}-\underline{\mu}_j)^T\Sigma(x-\mu_j)$$

$$\cancel{\underline{\underline{x}}^T\underline{\underline{\Sigma}}^{-1}\underline{x}} - \underbrace{2\underline{\mu}_i^T\underline{\underline{\Sigma}}^{-1}}\underline{x} + \underline{\mu}_i^T\underline{\underline{\Sigma}}\mu_i < \cancel{\underline{x}^T\underline{\underline{\Sigma}}^{-1}\underline{x}} - 2\mu_j^T\underline{\underline{\Sigma}}^{-1}\underline{x} + \underline{\mu}_j^T\underline{\underline{\Sigma}}^{-1}\underline{\mu}_j$$

$$\underbrace{2\mu_i^T\underline{\underline{\Sigma}}^{-1}}\underline{x} - \underbrace{\mu_i^T\Sigma\mu_i} > 2\mu_j^T\Sigma^{-1}\underline{x} - \underline{\mu}_j^T\underline{\Sigma}^{-1}\mu_j$$



CLASSIFICATORE LINEARE (AFFINE)

$$(\underline{x}-\underline{\mu}_i)^T\Sigma^{-1}(\underline{x}-\underline{\mu}_i) < (\underline{x}-\underline{\mu}_j)^T\Sigma^{-1}(x-\mu_j)$$

$$\underline{\underline{\Sigma}} = \underline{\underline{I}} \quad \underline{\underline{\Sigma}} = \sigma^2\underline{\underline{I}}_N \Rightarrow (\underline{x}-\underline{\mu}_i)^T(\underline{x}-\underline{\mu}_i) < (\underline{x}-\underline{\mu}_j)^T(\underline{x}-\underline{\mu}_j)$$

$$d^2(\underline{x},\underline{\mu}_i) < d^2(\underline{x},\underline{\mu}_j)$$

TESSELLAZIONE DI VORONOI

$$\underline{\underline{\Sigma}} = Q\Lambda Q^T \qquad \Sigma^{-1} = Q\Lambda^{-1}Q^T \qquad \Sigma\Sigma^{-1} = \underline{\underline{I}}$$

$$\Lambda^{-1/2}\Lambda^{-1/2}$$

$$(\underline{x}-\mu_i)^T\Sigma^{-1}(x-\mu_i) < (x-\mu_j)^T\Sigma^{-1}(x-\mu_j)$$

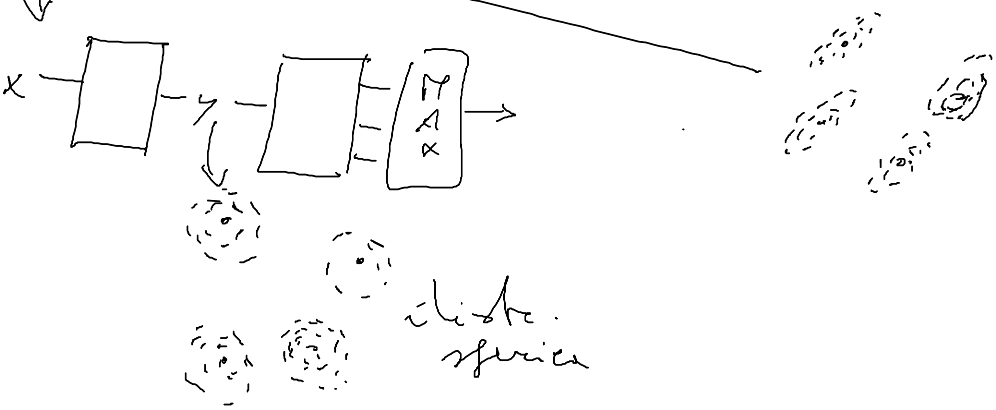
$$(\underline{x}-\underline{\mu}_i)^T \Sigma^{-1}(\underline{x}-\underline{\mu}_i) < (\underline{x}-\underline{\mu}_j)^T \Sigma^{-1}(\underline{x}-\underline{\mu}_j)$$

$$(x-\mu_i)^T Q\Lambda^{-1/2}\Lambda^{-1/2}Q^T(x-\mu_i) < \ldots$$

$$(x^T Q\Lambda^{-1/2} - \mu_i^T Q\Lambda^{-1/2})(\Lambda^{-1/2}Q^T x - \Lambda^{-1/2}Q^T \mu_i)$$

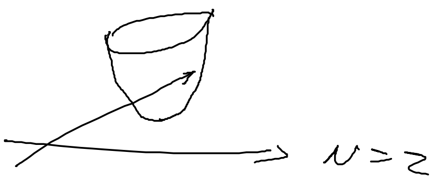
$$\boxed{\underline{y} = \underline{\underline{\Lambda}}^{-1/2}\underline{\underline{Q}}^T \underline{x}} \qquad \boxed{m_i = \Lambda^{-1/2}Q^T \underline{\mu}_i}$$

$$(y^T - m_i^T)(\underline{y} - \underline{m}_i) > (\underline{y}^T - m_j^T)(\underline{y} - \underline{m}_j)$$



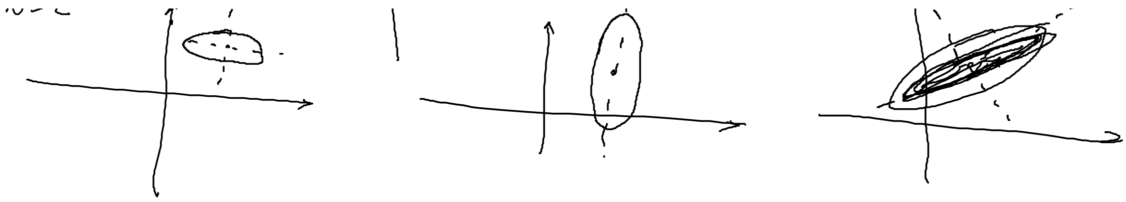
GAUSSIANA MULTIDIMENSIONALE

$$f_{\underline{x}}(\underline{x}) = N(\underline{x};\underline{\mu},\underline{\underline{\Sigma}}) = \frac{1}{(2\pi)^{N/2}|\Sigma|}\, e^{-\frac{1}{2}(\underline{x}-\underline{\mu})\Sigma^{-1}(\underline{x}-\underline{\mu})}$$

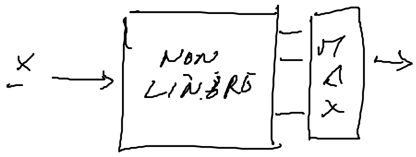


$$\underline{\underline{\Sigma}} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

N = 2

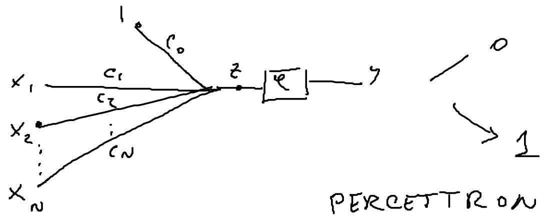


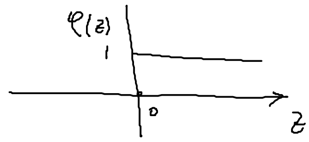
$$\frac{f_{\underline{x}}(\underline{x}|a_i)\,p_i}{\cancel{f_{x}(x)}} \overset{a_i}{>} \frac{f_{\underline{x}}(\underline{x}|a_j)\,p_j}{\cancel{f_{x}(x)}}$$



#1

$$\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_N \end{bmatrix}$$

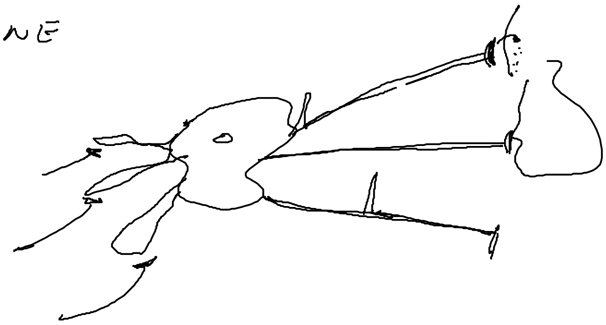


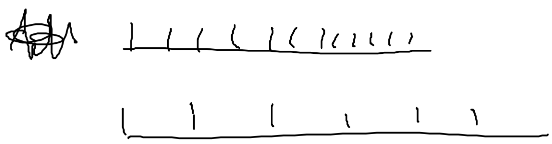


PERCETTRONE

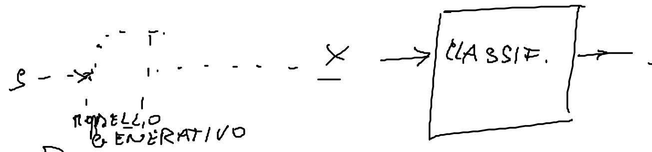
$$z = \varphi(c^T x + c_0)$$

$$\varphi(c^T x)$$





$\underline{x} \in \mathcal{X} \subseteq \mathbb{R}^N$



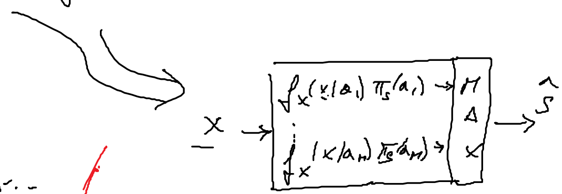
$\hat{s} \in \mathcal{A} = \{a_1, \dots a_M\}$

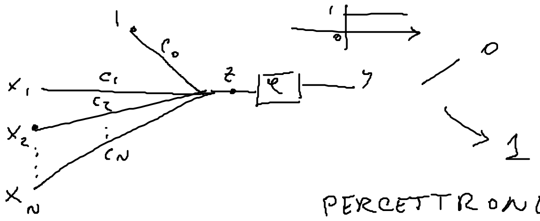
$s \in \mathcal{A}$

MODEL-BASED $\Rightarrow$ $f_{\underline{x}}(\underline{x}|s)$ $\pi_s(s)$ $\Rightarrow$ MAP $f_{\underline{x}}(\underline{x}|a_i)\pi_s(a_i) > f_{\underline{x}}(\underline{x}|a_j)\pi_s(a_j)$ $\forall j \neq i$

likelihood

DATA-DRIVEN





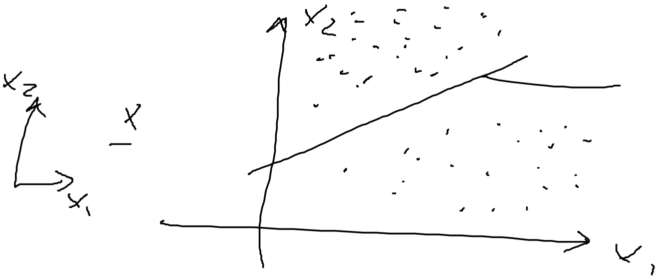
PERCETTRONE

CLASSIFICATORE BINARIO

$\mathcal{A} = \{a_1, a_2\}$

$\underline{x} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_N \end{bmatrix}$

$z = \underline{c}^T \underline{x}$

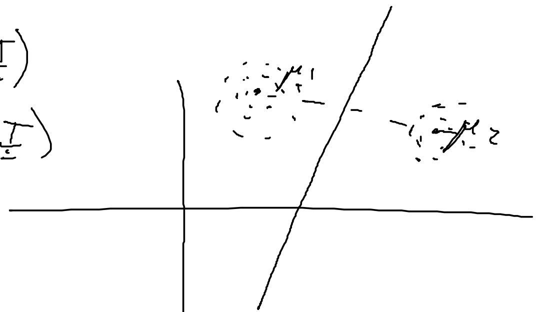


$c_1 x_1 + c_2 x_2 + c_0 = 0$

Se i dati sono gaussiani (condizionalmente gaussiani)

$f_{\underline{x}}(\underline{x}|a_1) = \mathcal{N}(\underline{x}; \underline{\mu}_1, \sigma^2 \underline{I})$

$f_{\underline{x}}(\underline{x}|a_2) = \mathcal{N}(\underline{x}, \mu_2, \sigma^2 \underline{I})$

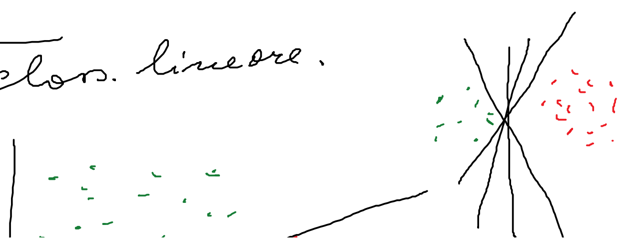


Se i dati sono arbitrari (data-driven)

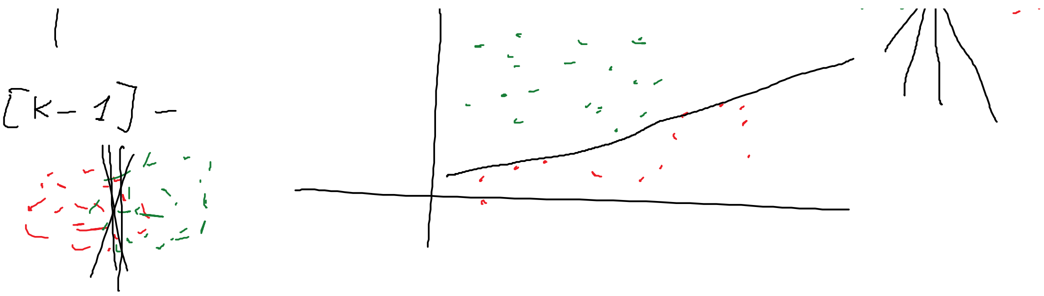
$\{(\underline{x}[n], s[n]),\ n = 1, \dots \ell\}$ TRAINING SET

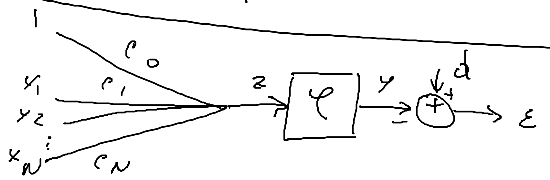
$s[n] \in \{0, 1\}$

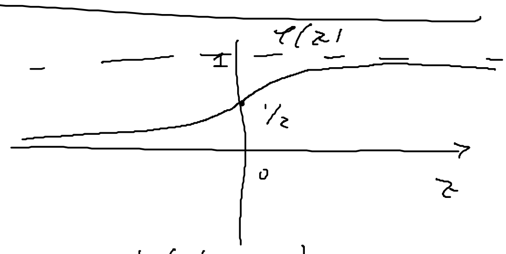
class. lineare



$$C[k] = c[k-1] -$$





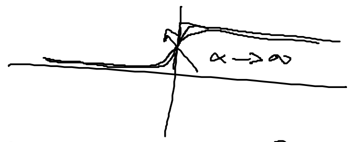


$$c[k] = c[k-1] - \eta \nabla_{\underline{c}} \mathcal{E}(\underline{c}(k-1))$$

~~BATCH~~ →

STOCASTICO

$$\varphi(z) = \frac{1}{1+e^{-\alpha z}} \qquad \varphi(z) = \frac{1}{1+e^{-z}}$$



$$\varphi'(z) = \varphi(z)\left(1-\varphi(z)\right)$$

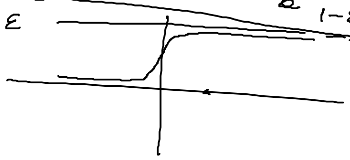
$$\mathcal{E}(c[k-1]) = \cancel{\sum_{n=1}^{\ell}} \left(d[\underset{\uparrow k}{n}] - \varphi\left(\underline{c}^T[k-1]\,\underline{x}[\underset{\uparrow k}{n}]\right)\right)^2$$

$$\nabla_{\underline{c}} \mathcal{E} = \nabla_{\underline{c}} \left(d - \varphi(c^T x)\right)^2 = -2\left(d-\varphi(c^T x)\right)\varphi'(c^T x)\,\underline{x}$$

$$= -2\left(d-\varphi(c^T x)\right)\underbrace{\varphi(c^T x)}_{y}\underbrace{\left(1-\varphi(c^T x)\right)}_{1-y}\underline{x}$$

$$c[k] = c[k-1] + \eta\,\underline{x}[k]\,\varphi(c^T[k-1]x[k])\left(1-\varphi(c^T[k-1]x[k])\right)\left(d[k] - \varphi(c^T[k-1]\,x[k])\right)$$

1-ε ε 1-ε

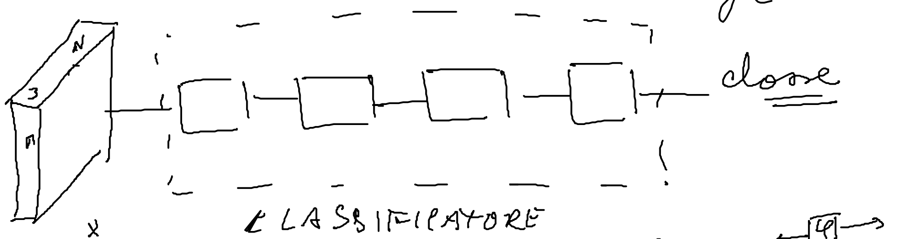


1-ε — 0

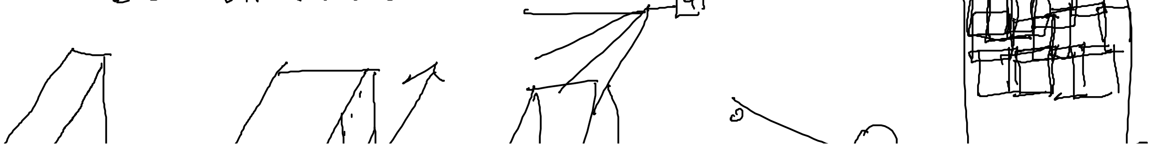
0 — (1-ε)

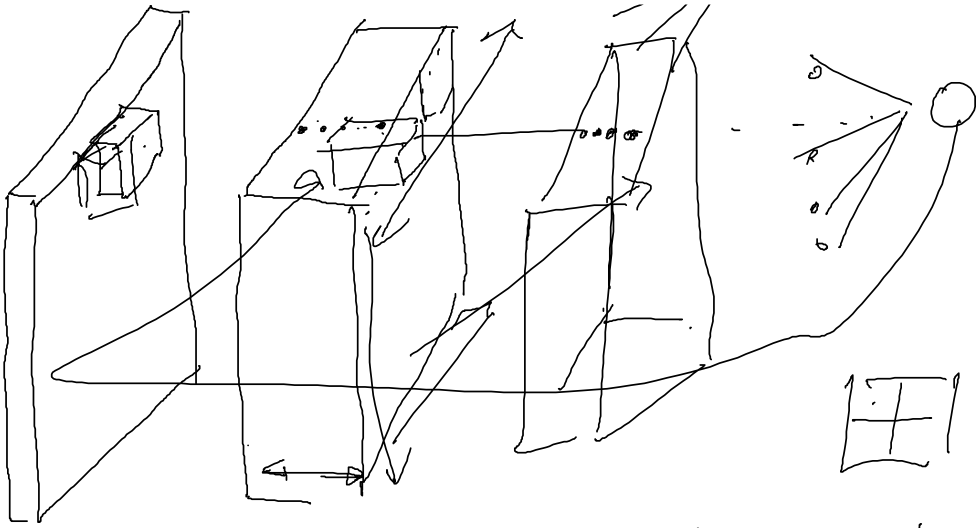
DATA-DRIVEN TRAINING SET TEST SET VALIDATION SET

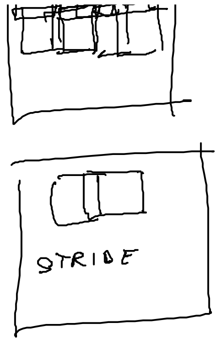
progett. sviluppo.



CLASSIFICATORE

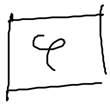


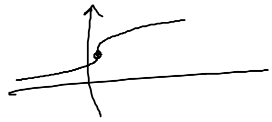




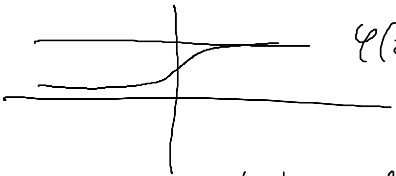
POOLING: $\left\{ \begin{array}{l} \text{MEDIA} \\ \text{MAX} \end{array} \right.$

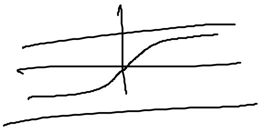
TIPI DI NON LINEARITA'





$$\varphi(z) = \frac{1}{1+e^{-z}} = \ell(z)$$



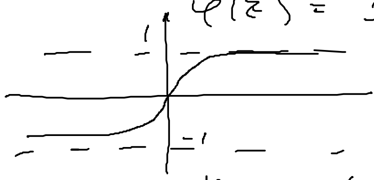


$$\varphi(z) = \frac{e^{z} - e^{-z}}{e^{z}+e^{-z}} = \operatorname{tgh}(z) \qquad \left(\ell(z) - \frac{1}{2}\right)\cdot 2 =$$

$$= 2\left(\frac{1}{1+e^{-z}} - \frac{1}{2}\right) = \cancel{2}\,\frac{2-1+e^{-z}}{\cancel{2}\,(1+e^{-z})} = \frac{1-e^{-z}}{1+e^{-z}}$$

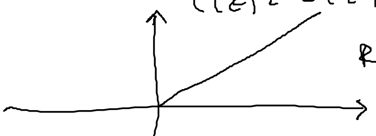
$$= \frac{e^{-z/2}}{e^{-z/2}} \qquad \frac{e^{\frac{z}{2}} - e^{-\frac{z}{2}}}{e^{\frac{z}{2}} + e^{-\frac{z}{2}}} = \operatorname{tgh}\left(\frac{z}{2}\right)$$

$\varphi(z) = r(z)$



RELU

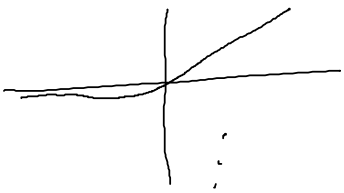
RECTIFICATION LOGIC UNIT



$\vdots$

SOFTMAX

$$0 \le y_i \le 1$$

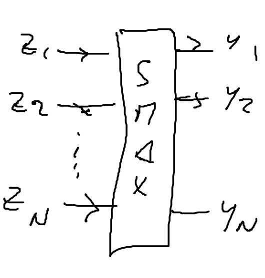


$$y_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}$$

$$i = 1, \dots, N$$

$$[1 \;\; 0 \;\; 3] = \left[\frac{e^{1}}{\cdot}, \frac{e^{2}}{\cdot}, \frac{e^{3}}{\cdot}\right]$$

$$\sum_{i=1}^{N} e^{z_i} = e^{1} + e^{0} + e^{3} =$$



MODEL-BASED CLASSIFIER

$s \dashrightarrow \cdots \to \underline{x} \to \boxed{\text{CLAS.}} \to \hat{s}$

$s \in \mathcal{A} = \{a_1, \dots a_M\}$
$\hat{s} \in \mathcal{A} = \{a_1, \dots a_M\}$

$$\hat{s}_o = \underset{\hat{s}}{\arg\min}\, P(e)$$

$$P(e) = \Pr\{\hat{s} \neq s\} \longrightarrow$$

$$\left.\begin{array}{l} p(s=a_1|\underline{x}) \\ p(s=a_2|\underline{x}) \\ \vdots \\ p(s=a_M|\underline{x}) \end{array}\right\} \longrightarrow \text{MAX } \hat{s}$$

$\min P(e) \longleftrightarrow \max p(s|\underline{x})$

Likelihood — prior — Verosimiglianza

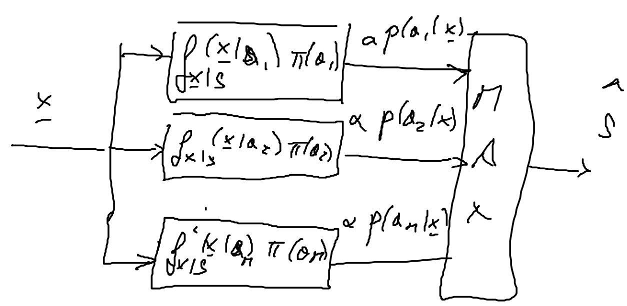
$$p(s|x) = \frac{f_{x|s}(\underline{x}|s)\,\pi(s)}{f_x(x)}$$

MAP (MASSIMA PROB. A POSTERIORI)

$$\hat{s} = a_i \quad \text{se} \quad p(a_i|\underline{x}) > p(a_j|\underline{x}) \quad \forall j \neq i$$

$$\frac{f_{x|s}(\underline{x}|a_i)\,\pi(a_i)}{\cancel{f_x(x)}} > \frac{f_{x|s}(x|a_j)\,\pi(a_j)}{\cancel{f_x(x)}}$$

$$\hat{s} = a_i \quad \text{se} \quad f_{\underline{x}|s}(\underline{x}|a_i)\,\pi(a_i) > f_{\underline{x}|s}(\underline{x}|a_j)\,\pi(a_j) \quad \forall j \neq i$$



$\Downarrow$

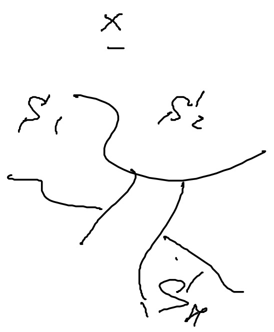
$$\log f_{\underline{x}|s}(\underline{x}|a_i) + \log \pi(a_i) > \log f_{\underline{x}|s}(\underline{x}|a_j) + \log \pi(a_j)$$

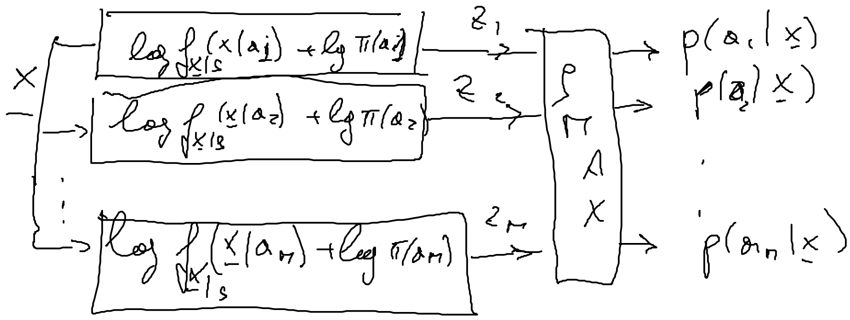
CALCOLO DELLE PROB. a posteriori

$$p(a_i|\underline{x}) = \frac{f_{x|s}(\underline{x}|a_i)\,\pi(a_i)}{\sum_{j=1}^{M} f_{x|s}(\underline{x}|a_j)\,\pi(a_j)} = \frac{e^{\overbrace{\log f_{x|s}(\underline{x}|a_i) + \log \pi(a_i)}^{z_i}}}{\sum_{j=1}^{M} e^{\overbrace{\log f_{x|s}(\underline{x}|a_j) + \log \pi(a_j)}^{z_j}}}$$

z_i

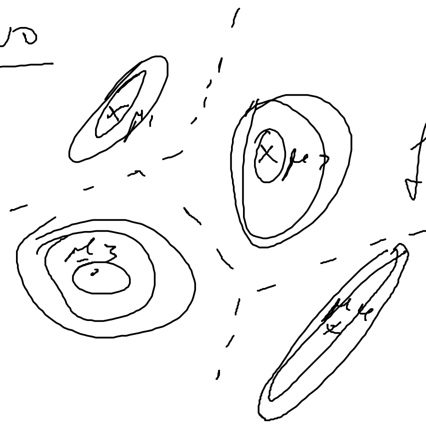
$$= \frac{e}{\sum_{j=1}^{M} e^{z_j}} \qquad \text{SOFTMAX}$$

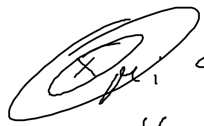




CASO GAUSSIANO

Σ_i diverse





$$f_{\underline{x}|s}(\underline{x}|a_i) = \frac{1}{(2\pi)^{N/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T \Sigma_i^{-1}(\underline{x}-\underline{\mu}_i)}$$

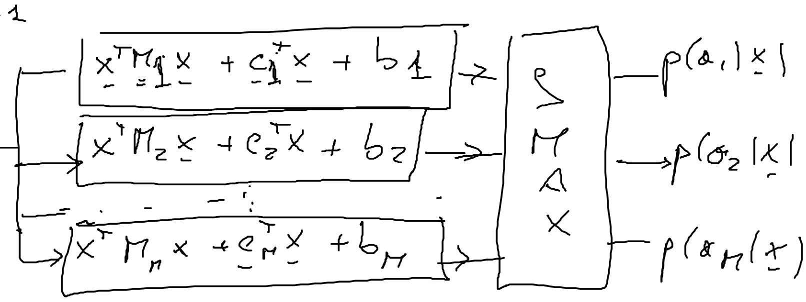
$$\log f_{\underline{x}|s}(\underline{x}|a_i) + \log \pi(a_i)$$

$$= \log \frac{1}{(2\pi)^{N/2}|\Sigma_i|} - \frac{1}{2}(\underline{x}-\underline{\mu}_i)^T \Sigma_i^{-1}(\underline{x}-\underline{\mu}_i) + \log \pi(a_i)$$

$$= \underbrace{-\frac{1}{2} x^T \Sigma_i^{-1} \underline{x}}_{\text{quadratica}} + \underbrace{\underline{\mu}_i^T \Sigma_i^{-1} \underline{x}}_{\text{lineare}} + \underbrace{\frac{1}{2}\mu_i^T \Sigma_i^{-1} \mu_i + \log \pi(a_i) - \log (2\pi)^{N/2} |\Sigma_i|^{1/2}}_{b_i}$$

$M_i := -\frac{1}{2}\Sigma_i^{-1}$

$\underline{c}_i = \Sigma_i^{-1}\mu_i$

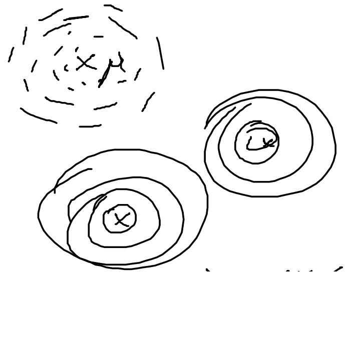


Caso gaussiano sferico

$$\Sigma_i = \Sigma = \sigma^2 I_N$$

$$\Sigma^{-1} = \frac{1}{\sigma^2} I_N$$

$$\log f_{x|s}(x|a_i) + \log \pi(a_i) =$$



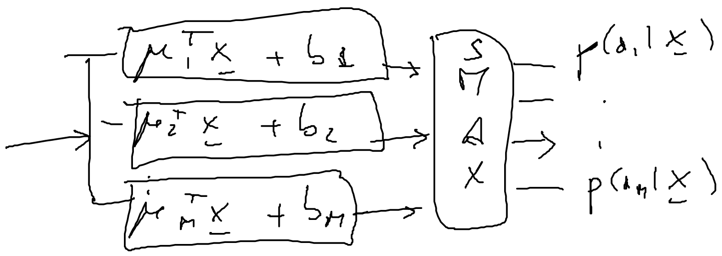
$$\log f_{x|s}(\underline{x}|a_i) + \log \pi(a_i) =$$

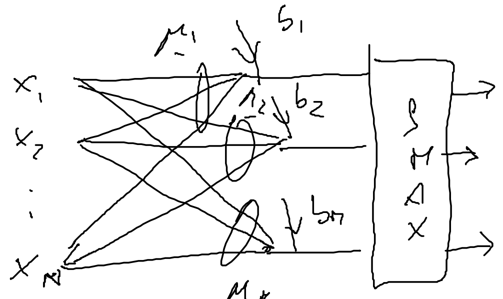
$$c - \frac{1}{\sigma^2} \ldots N$$

$$= \underbrace{\cancel{-\frac{1}{2} \underline{x}^T \underline{\underline{\Sigma}}^{-1} \underline{x}}}_{\text{quadratica}} + \underbrace{\underline{\mu}_i^T \underline{\underline{\Sigma}}^{-1} \underline{x}}_{\text{lineare}} + \underbrace{\frac{1}{2} \mu_i^T \Sigma^{-1} \mu_i + \log \pi(a_i) - \cancel{\log (2\pi)^{N/2} |\Sigma|^{1/2}}}$$

$$= \frac{1}{\sigma^2} \mu_i^T \underline{x} + \underbrace{\frac{1}{2\sigma^2} \mu_i^T \mu_i + \log \pi(a_i)}_{b_i}$$

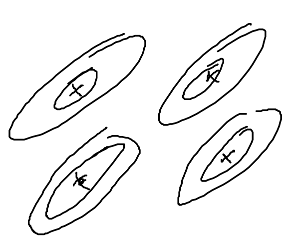
$$\frac{e^{z_i + c}}{\sum_{j=1}^{M} e^{z_j + c}} = \frac{\cancel{e^c}\, e^{z_i}}{\cancel{e^c} \sum e^{z_j}}$$





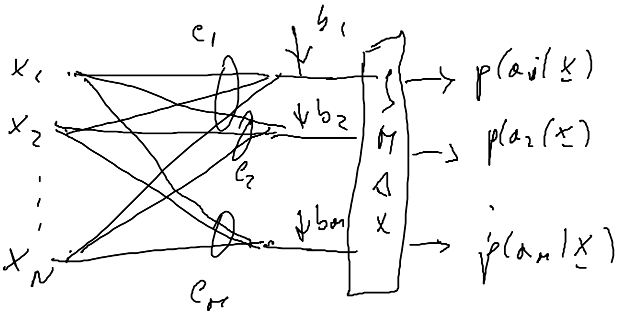
Gaussiane aventi la stessa covarianza

$$\underline{\underline{\Sigma}}_i = \underline{\underline{\Sigma}}$$



$$\log f_{x|s}(\underline{x}|a_i) + \log \pi(a_i)$$

$$= \underbrace{\cancel{-\frac{1}{2} \underline{x}^T \underline{\underline{\Sigma}}^{-1} \underline{x}}}_{\text{quadratica}} + \underbrace{\overbrace{\underline{\mu}_i^T \underline{\underline{\Sigma}}^{-1}}^{c_i^T} \underline{x}}_{\text{lineare}} + \underbrace{\frac{1}{2} \mu_i^T \Sigma^{-1} \mu_i + \log \pi(a_i) - \cancel{\log (2\pi)^{N/2} |\Sigma|^{1/2}}}_{b_i}$$



$$\Sigma \Sigma^{-1} = Q \cancel{\Lambda} \cancel{Q^T Q} \cancel{\Lambda^{-1}} Q^T$$

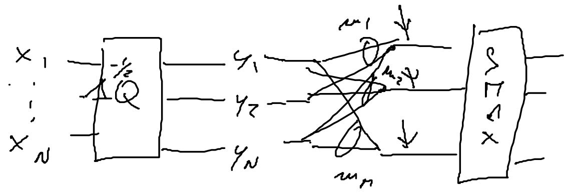
$$\underline{\underline{\Sigma}} = Q \Lambda Q^T \qquad \underline{\underline{\Sigma}}^{-1} = Q \Lambda^{-1} Q^T$$

$$\underline{\underline{Q}}^T \underline{\underline{Q}} = \underline{\underline{I}}_N$$

$$\log f_{\underline{x}|s}(\underline{x}|a_i) + \log \pi(a_i) = \underline{\mu}_i^T \underline{\underline{\Sigma}}^{-1} \underline{x} + \frac{1}{2} \mu_i^T \Sigma^{-1} \mu_i + \log \pi(a_i)$$

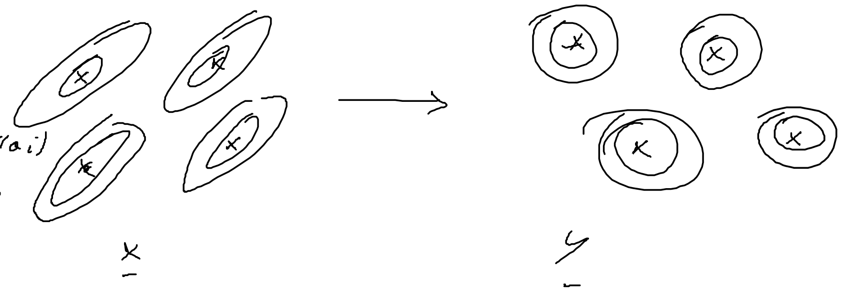
$$= \mu_i^T Q \Lambda^{-1} \underbrace{Q^T x} + \frac{1}{2} \mu_i^T Q \Lambda^{-1} Q^T \mu_i + \log \pi(a_i) \qquad y = \Lambda^{-\frac{1}{2}} Q^T x$$

$$m_i = \Lambda^{-\frac{1}{2}} \underline{\underline{Q}}^T \underline{\mu}_i$$

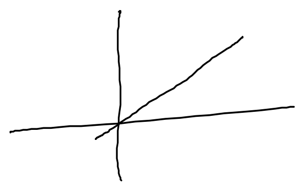


$$\mathrm{cov}[y] = \underline{\underline{S}}_i = I_N$$

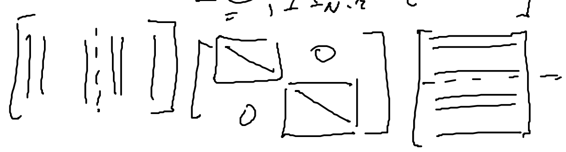
$$\underbrace{\underline{m}_i^T \underline{y} + \frac{1}{2} m_i^T m_i + \log \pi(a_i)}$$



RANK-DEFICIENT $\quad r < N \qquad N = 2$

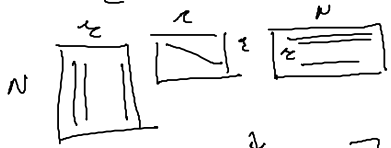


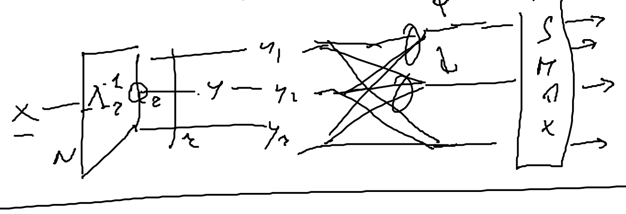
$$\underline{\underline{\Sigma}} = \underset{(N)}{\underline{\underline{Q}}\,\underline{\underline{\Lambda}}\,\underline{\underline{Q}}^T} = [Q_r \mid Q_{N-r}] \left[\begin{array}{c|c} \Lambda_r & \underline{\underline{0}} \\ \hline \underline{\underline{0}} & \Lambda_{N-r} \end{array}\right] \begin{bmatrix} Q_r^T \\ Q_{N-r}^T \end{bmatrix}$$



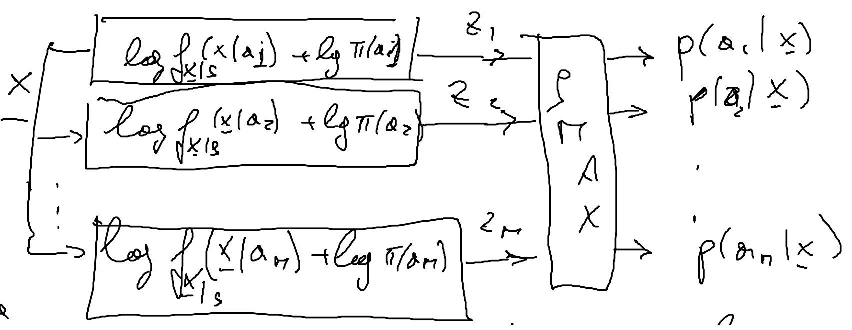
$$\Lambda_{N-r} = \underline{\underline{0}}$$

$$= Q_r \Lambda_r Q_r^T \qquad \underline{y} =$$





Altre distribuzioni oltre alla Gaussiana



$\hookrightarrow \boxed{\log f_{\underline{x}|s}(\underline{x}|a_M) + \log \pi(a_M)} \longrightarrow \boxed{\ } \to p(a_n|\underline{x})$

Funzioni che forniscono un classificatore semplice, per esempio lineare?

CLASSE ESPONENZIALE

$$\boxed{f_{\underline{x}|s}(\underline{x}|a_i) = \underbrace{\hat{h}(\underline{x})}_{\text{non dipende da } i}\, e^{\underbrace{\eta_i^T \underline{T}(\underline{x}) - c_i}_{\text{dipende da } i}}}$$

Per essere una valida distribuzione

$$\int_{\mathcal{X}} f_{\underline{x}|s}(\underline{x}|a_i)\, d\underline{x} = 1 \qquad \forall i$$

$$\int_{\mathcal{X}} h(\underline{x})\, e^{\eta_i^T \underline{T}(\underline{x}) - c_i}\, dx = 1$$

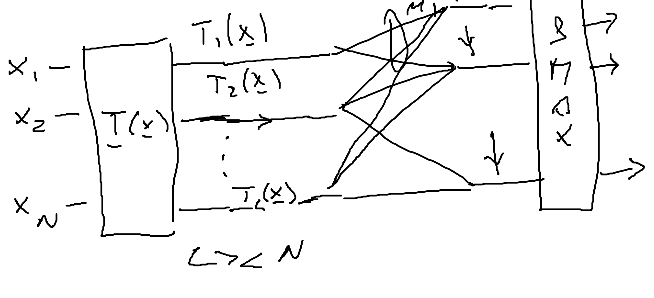
$$\int_{\mathcal{X}} h(\underline{x})\, e^{\eta_i^T \underline{T}(\underline{x})}\, dx = e^{c_i}$$

FUNZIONE DI PARTIZIONE

$$c_i = \log \int_{\mathcal{X}} h(\underline{x})\, e^{\eta_i^T \underline{T}(\underline{x})}\, dx$$

$f_{\underline{x}|s}(\underline{x}|a_i)\,\pi(a_i)$ oppure $\log f_{\underline{x}|s}(\underline{x}|a_i) + \log \pi(a_i) =$

$$= \underbrace{\cancel{\log h(\underline{x})}}_{\text{non dipende da } i} + \eta_i^T \underline{T}(\underline{x}) \underbrace{- c_i + \log \pi(a_i)}$$



La funzione appartiene alla famiglia esponenziale?

$$f_{\underline{x}|s}(\underline{x}|a_i) = \frac{1}{(2\pi)^{N/2}|\underline{\Sigma}|}\, e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T \underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_i)} \overset{?}{=} h(\underline{x})\, e^{\eta_i^T T(\underline{x}) - c_i}$$

$$= \frac{1}{(2\pi)^{N/2}|\Sigma|}\, e^{-\frac{1}{2}x^T \underline{\Sigma}^{-1} x}\, e^{\overbrace{\underbrace{\mu_i^T \underline{\Sigma}^{-1}}_{\eta_i^T} \underbrace{\underline{x}}_{\underline{x}}}^{\eta_i \ \underline{T}(\underline{x})} - \underbrace{\frac{1}{2}\mu_i^T \mu_i}_{c_i}}$$

$$f_{\underline{x}|s} \quad \ldots \quad i=1$$

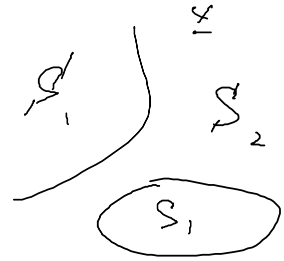
$$= \underbrace{\frac{1}{(2\pi)^{N/2}|\Sigma|} e^{-\frac{1}{2}x^T \underline{\underline{\Sigma}}^{-1} x}}_{h(x)} \; e^{\underbrace{\mu_i \Sigma^{-1} \underline{x}}_{\mu_i^T \;\; \underline{x}} - \underbrace{\frac{1}{2}\mu_i \mu_i}_{c_i}}$$

$$\underline{T}(\underline{x}) = x$$

$$\rightarrow \underline{T}(\underline{x}) = \underline{\underline{\Sigma}}^{-1} \underline{x}$$

CASO BINARIO

$$\Omega = \{a_1 \; a_2\}$$



$$\underline{x} \rightarrow \boxed{} \rightarrow \hat{s} \qquad \begin{cases} p(s=a_1|\underline{x}) \\ p(s=a_2|x) = 1 - p(s=a_1|x) \end{cases}$$

MAP

$$f_{x|s}(\underline{x}|a_1)\,\pi(a_1) \underset{a_2}{\overset{a_1}{\gtrless}} f_{\underline{x}|s}(\underline{x}|a_2)\,\pi(a_2)$$

RADAR

target present/not present

bit comunicazioni

bit $\begin{matrix} \nearrow 0 \\ \searrow 1 \end{matrix}$

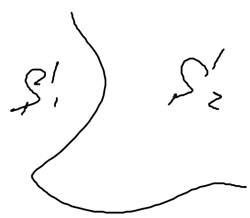
sist allarme $\begin{matrix} \nearrow \text{SI} \\ \searrow \text{NO} \end{matrix}$

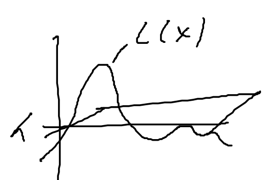
$$\frac{f_{x|s}(x|a_1)\,\pi(a_1)}{f_{x|s}(x|a_2)\,\pi(a_2)} \underset{a_2}{\overset{a_1}{\gtrless}} 1$$

$$\frac{f_{x|s}(\underline{x}|a_1)}{f_{\underline{x}|s}(x|a_2)} \underset{a_2}{\overset{a_1}{\gtrless}} \frac{\pi(a_2)}{\pi(a_1)}$$

$$\underline{x} \rightarrow \boxed{L(\underline{x})} \rightarrow \underset{\lambda}{\boxed{\begin{matrix}\text{rivela} \\ \text{soglia}\end{matrix}}} \rightarrow \begin{cases} >\lambda & a_1 \\ <\lambda & a_2 \end{cases}$$

$$\underbrace{\boxed{\log \frac{f_{x|s}(\underline{x}|a_1)}{f_{x|s}(\underline{x}|a_2)}}}_{l(\underline{x})} \underset{a_2}{\overset{a_1}{\gtrless}} \underbrace{\log \frac{\pi(a_2)}{\pi(a_1)}}_{\lambda}$$





$$\times \mu_1 \qquad \underset{\times}{\mu_2}$$

$$\log\left[\frac{\frac{1}{(2\pi)^{N/2}|\Sigma_1|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\mu_1)^T \underline{\underline{\Sigma}}_1^{-1}(\underline{x}-\underline{\mu}_1)}}{\frac{1}{(2\pi)^{N/2}|\Sigma_2|^{1/2}} e^{-\frac{1}{2}(x-\mu_2)^T \Sigma_2^{-1}(\underline{x}-\mu_2)}}\right]$$

$$-\underbrace{\log(2\pi)^{N/2}|\Sigma_1|^{1/2} + \log(2\pi)^{N/2}|\Sigma_2|^{1/2}} - \frac{1}{2}(\underline{x}-\underline{\mu}_1)^T\underline{\underline{\Sigma}}_1^{-1}(\underline{x}-\underline{\mu}_1) + \frac{1}{2}(\underline{x}-\underline{\mu}_2)^T\underline{\underline{\Sigma}}_2^{-1}(\underline{x}-\underline{\mu}_2)$$

$$= -\frac{1}{2}\left(\underline{x}^T\Sigma_1^{-1}\underline{x} - \underline{x}^T\underline{\underline{\Sigma}}_2^{-1}\underline{x} + \underbrace{\mu_1^T\Sigma_1^{-1}\mu_1 - \mu_2^T\Sigma_2^{-1}\mu_2}\right) + \mu_1^T\Sigma_1^{-1}\underline{x} - \mu_2^T\Sigma_2^{-1}\underline{x}$$

$$\text{cost} + \underline{x}^T\left(\Sigma_1^{-1} - \Sigma_2^{-1}\right)\underline{x} + \text{cost}' + \underbrace{\left(\mu_1^T\Sigma_1^{-1} - \mu_2^T\Sigma_2^{-1}\right)}_{c^T}\underline{x}$$

Lezione 23

Lunedi 13 Dicembre 14-16

Introduzione all'uso del software Tensorflow per l'addestramento di reti neurali
Tenuta dall'Ing. Giovanni Di Gennaro

Lezione 24

mercoledì 15 dicembre 2021

9:11

Prima parte: Classificatore gaussiano, caso Binario; Curve ROC

Seconda parte: Amedeo Buonanno sulle applicazioni del Mchine Learning ai sistemi energetici