

TRASFORMATE DECORRELANTI DISCRETE

lezioni del corso di
TRASMISSIONE ED ELABORAZIONE NUMERICA
DEI SEGNALE

AA. 2015-16

Prof. FRANCESCO A.N. FALIERI

Dato un vettore N -dimensionale stocastico $\underline{x}$, poniamo il problema di trovare una matrice quadrata $\underline{C}$ tale che $\underline{X} = \underline{C}\underline{x}$ sia a componenti non correlate, ovvero che la matrice di covarianza

$$\text{cov}[\underline{X}] = E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^H] = \text{diagonale}.$$

Per semplificare l'analisi che segue ometteremo che il vettore $\underline{x}$ non a media nulla $E[\underline{x}] = \underline{0}$, pertanto $E[\underline{X}] = \underline{C}E[\underline{x}] = \underline{0}$ e poniamo studiare il problema con riferimento alle matrici di correlazione $E[\underline{x}\underline{x}^H] = E[\underline{X}\underline{X}^H]$.

NOTA 1 Per evidenziare che questa omissione non ci fa perdere in generalità, immaginiamo che $E[\underline{x}] = \underline{\mu}$. In tal caso poniamo definire un nuovo vettore $\underline{y} = \underline{x} - \underline{\mu}$ a media nulla riprendendo la analisi visto che

$$\begin{aligned} \text{cov}[\underline{X}] &= \text{cov}[\underline{C}\underline{x}] = E[(\underline{X} - \underline{C}\underline{\mu})(\underline{X} - \underline{C}\underline{\mu})^H] \\ &= \underline{R}_x - \underline{C}\underline{\mu}\underline{\mu}^H\underline{C}^H = \underline{R}_y \end{aligned}$$

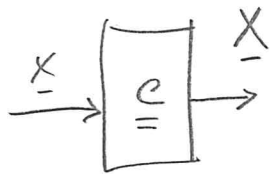
Il vettore $\underline{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix}$ è in generale complesso, così come la trasformata $\underline{C}$. Quindi anche $\underline{X} = \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix}$ è complesso.

Il vettore $\underline{x}$ può provenire da una serie temporale TDZ

$$\dots \underline{x}[n] \underline{x}[n+1] \underline{x}[n+2] \dots$$

$\underbrace{\hspace{10em}}_N$
 $\downarrow$
 $\underline{x}$

o da un qualunque altro processo stocastico stazionario.
 Rappresentiamo la trasformazione con un blocco $\underline{C}$ (quadrato per ora)



e poniamo il problema di trovare $\underline{C}$ tale che

$$E[\underline{x}\underline{x}^H] = \text{diagonale}.$$

Notiamo che per qualunque $\underline{C}$, la matrice di auto correlazione dell'uscita è

$$E[\underline{x}\underline{x}^H] = E[\underline{C}\underline{x}\underline{x}^H\underline{C}^H] = \underline{C} E[\underline{x}\underline{x}^H] \underline{C}^H = \underline{C} \underline{R}_x \underline{C}^H$$

$\boxed{N} \quad \boxed{N} \quad \boxed{N}$

Supponiamo ora che $\underline{R}_x$ sia fattorizzabile

$$\underline{R}_x = \underline{F} \underline{F}^H$$

$\boxed{N} \quad \boxed{N}$

e che il "fattore" $\underline{F}$ (detta radice quadrata di $\underline{R}_x$) sia invertibile.

PROPOSIZIONE: Per $\underline{C} = \alpha \underline{F}^{-1}$, dove α è una qualunque costante, $E[\underline{x}\underline{x}^H] = \text{diagonale}$

PROVA :

$$E[\underline{x}\underline{x}^H] = \underline{C} \underline{R}_x \underline{C}^H = \alpha \underline{F}^{-1} \underline{F} \underline{F}^H \underline{F}^{-H} \alpha = \alpha^2 \underline{I}_N$$

Anche la fattorizzazione è del tipo

TD3

$$\underline{R}_x = \underline{F} \underline{\Lambda} \underline{F}^H \quad \text{con } \underline{\Lambda} \text{ diagonale,}$$

$$\begin{matrix} N & N & N \\ \square & \square & \square \end{matrix}$$

$$E[\underline{X}\underline{X}^H] \underline{F}^{-1} \underline{F} \underline{\Lambda} \underline{F}^H \underline{F}^{-H} = \underline{\Lambda} \quad \text{diagonale}$$

La fattorizzazione di una matrice $\underline{R}_x$ (Hermitiana) definita
 positiva, non è unica. Pertanto vanno imposti ulteriori
 vincoli per trovare l'elemento la trasformazione che
 viene della fattorizzazione applicata a segnali di rumore è anche
FATTORIZZAZIONE DI CHOLESKY detta FILTRO A RADICE QUADRATA

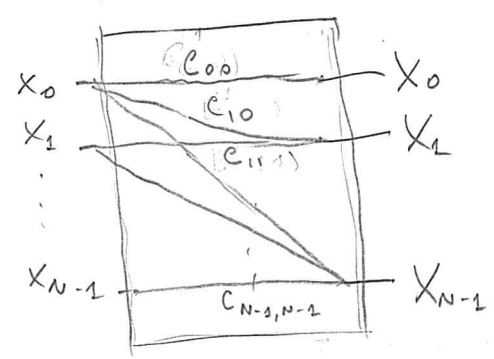
Se si richiede che il fattore $\underline{F}$ sia triangolare
 inferiore, la fattorizzazione è unica (dell'algebra lineare)

$$\underline{R}_x = \underline{G} \underline{G}^H$$

$$\begin{matrix} N & N \\ \triangle & \triangle \end{matrix}$$

Pertanto $\underline{F} = \underline{G}$ e $\underline{C} = \underline{F}^{-1} = \underline{G}^{-1}$ (ancora triangolare più
 l'inverso di una matrice triangolare è triangolare (vedi nota 4))

Profondamente



$$\underline{G} = \begin{pmatrix} g_{00} & 0 & \dots & 0 & 0 \\ g_{10} & g_{11} & & & \\ g_{20} & g_{21} & g_{22} & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{N-1,0} & \dots & \dots & g_{N-1,N-1} \end{pmatrix}$$

$$\underline{C} = \begin{pmatrix} c_{00} & 0 & \dots & 0 & 0 \\ c_{10} & c_{11} & & & \\ c_{20} & c_{21} & c_{22} & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N-1,0} & c_{N-1,1} & \dots & c_{N-1,N-1} \end{pmatrix}$$

$$\underline{C} = \underline{G}^{-1}; \quad \underline{X} = \underline{G}^{-1} \underline{X}$$

$$\begin{matrix} \parallel & \triangle & \parallel \end{matrix}$$

Quindi $E[\underline{X}\underline{X}^H] = \underline{I}_N$
 La trasformazione è anche denominata
TRASFORMATA TRIANGOLARE IN FORMA DIRETTA

Per convincersi che l'inversa di una matrice triangolare è triangolare
 Consideriamo

$$\begin{pmatrix} g_{00} & 0 & 0 \\ g_{10} & g_{11} & 0 \\ g_{20} & g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} c_{00} & c_{01} & c_{02} \\ c_{10} & c_{11} & c_{12} \\ c_{20} & c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$N=3$$

Regole per colonne:

$$g_{00} c_{00} = 1 ; \quad g_{00} c_{01} = 0 ; \quad g_{00} c_{02} = 0 \quad \Rightarrow \quad c_{01} = 0 \quad c_{02} = 0$$

$$g_{10} c_{00} + g_{11} c_{10} = 0 ; \quad g_{10} c_{01} + g_{11} c_{11} = 1 ; \quad g_{10} c_{02} + g_{11} c_{12} = 0 \quad \Rightarrow \quad c_{12} = 0$$

$$g_{20} c_{00} + g_{21} c_{10} + g_{22} c_{20} = 0 ; \quad g_{20} c_{01} + g_{21} c_{11} + g_{22} c_{21} = 0$$

$$g_{20} c_{02} + g_{21} c_{12} + g_{22} c_{22} = 1$$

Facilmente estendibile a qualunque N. Ora resta la struttura dell'inversa in colonne facilmente in maniera iterativa

$$\begin{pmatrix} g_{00} & 0 & 0 \\ g_{10} & g_{11} & 0 \\ g_{20} & g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} c_{00} & 0 & 0 \\ c_{10} & c_{11} & 0 \\ c_{20} & c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Procediamo per colonne

$$\begin{cases} g_{00} c_{00} = 1 & \Rightarrow & c_{00} = \frac{1}{g_{00}} \\ g_{10} c_{00} + g_{11} c_{10} = 0 & \Rightarrow & c_{10} = \frac{1}{g_{11}} (-g_{10} c_{00}) \\ g_{20} c_{00} + g_{21} c_{10} + g_{22} c_{20} = 0 & \Rightarrow & c_{20} = \frac{1}{g_{22}} (-g_{20} c_{00} - g_{21} c_{10}) \end{cases}$$

$$\begin{cases} g_{11} c_{11} = 1 & \Rightarrow & c_{11} = \frac{1}{g_{11}} \\ g_{21} c_{11} + g_{22} c_{21} = 0 & \Rightarrow & c_{21} = \frac{1}{g_{22}} (-g_{21} c_{11}) \end{cases}$$

$$\begin{cases} g_{22} c_{22} = 1 & \Rightarrow & c_{22} = \frac{1}{g_{22}} \end{cases}$$

In generale

TD 5

$$\begin{pmatrix} g_{00} & 0 & 0 & \dots & 0 \\ g_{10} & g_{11} & 0 & \dots & 0 \\ g_{20} & g_{21} & g_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{N-1,0} & g_{N-1,1} & g_{N-1,2} & \dots & g_{N-1,N-1} \end{pmatrix} \begin{pmatrix} c_{00} & 0 & 0 & \dots & 0 \\ c_{10} & c_{11} & 0 & \dots & 0 \\ c_{20} & c_{21} & c_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N-1,0} & c_{N-1,1} & c_{N-1,2} & \dots & c_{N-1,N-1} \end{pmatrix} = \mathbb{I}_N$$

I^a colonne

$$\begin{cases} g_{00} c_{00} = 1 \\ g_{10} c_{00} + g_{11} c_{10} = 0 \\ g_{20} c_{00} + g_{21} c_{10} + g_{22} c_{20} = 0 \\ \vdots \\ g_{N-1,0} c_{00} + g_{N-1,1} c_{10} + \dots + g_{N-1,N-1} c_{N-1,0} = 0 \end{cases} \Rightarrow \begin{cases} c_{00} = \frac{1}{g_{00}} \\ c_{10} = \frac{1}{g_{11}} (-g_{10} c_{00}) \\ c_{20} = \frac{1}{g_{22}} (-g_{20} c_{00} - g_{21} c_{10}) \\ \vdots \\ c_{N-1,0} = \frac{1}{g_{N-1,N-1}} (-g_{N-1,0} c_{00} - g_{N-1,1} c_{10} - \dots - g_{N-1,N-2} c_{N-2,0}) \end{cases}$$

I^a colonne di C

$$\begin{cases} c_{00} = \frac{1}{g_{00}} \\ c_{i0} = -\frac{1}{g_{ii}} \sum_{l=0}^{i-1} g_{il} c_{l0} \quad i=1, \dots, N-1 \end{cases}$$

II^a colonne

$$\begin{cases} g_{11} c_{11} = 1 \\ g_{21} c_{11} + g_{22} c_{21} = 0 \\ \vdots \\ g_{N-1,1} c_{11} + g_{N-1,2} c_{21} + \dots + g_{N-1,N-1} c_{N-1,1} = 0 \end{cases}$$

$$c_{11} = \frac{1}{g_{11}}$$

$$c_{21} = \frac{1}{g_{22}} (-g_{21} c_{11})$$

$$\vdots$$

$$c_{N-1,1} = \frac{1}{g_{N-1,N-1}} (-g_{N-1,1} c_{11} - g_{N-1,2} c_{21} - \dots - g_{N-1,N-2} c_{N-2,1})$$

II^a colonne di C

$$\begin{cases} c_{11} = \frac{1}{g_{11}} \\ c_{i1} = -\frac{1}{g_{ii}} \sum_{l=1}^{i-1} g_{il} c_{l1} \quad i=2, \dots, N-1 \end{cases}$$

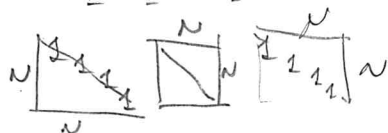
una colonna di C

$$\begin{cases} C_{jj} = \frac{1}{g_{jj}} \\ C_{ij} = -\frac{1}{g_{ii}} \sum_{l=j}^{i-1} g_{il} C_{lj} \end{cases} \quad i = j+1, \dots, N-1$$

↓
fin all'ultima colonna (rel $C_{N-1, N-1}$)

La fattorizzazione di Cholesky può essere organizzata facendo comparire tutti 1 sulle diagonali del fattore, includendo al centro una matrice diagonale Λ .

$$R_X = L \Lambda L^H$$

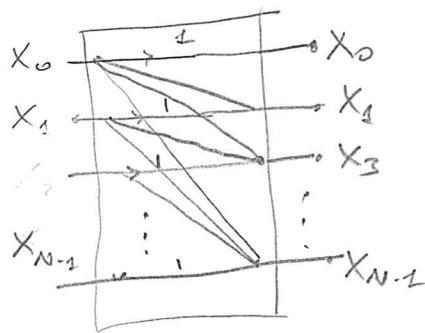


$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{10} & 1 & 0 & 0 \\ l_{20} & l_{21} & 1 & 0 \\ l_{30} & l_{31} & l_{32} & 1 \end{pmatrix} \quad N=4$$

$F = L$, $C = F^{-1} = L^{-1}$ anche questa triangolare con tutti 1 sulle diagonali.

$$\underline{X} = \underline{L}^{-1} \underline{X}$$

$$L^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ c_{10} & 1 & 0 & 0 \\ c_{20} & c_{21} & 1 & 0 \\ c_{30} & c_{31} & c_{32} & 1 \end{pmatrix}$$



FORMA
DIRETTA

$$E[\underline{X}\underline{X}^H] = \underline{\Lambda}$$

Osservando che

TD7

$$\underline{L} \underline{X} = \underline{x}$$

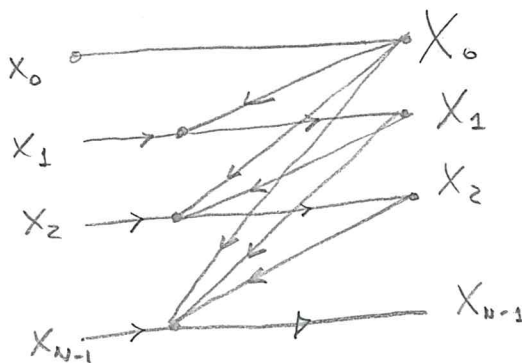
$\underline{X}$ si può calcolare per "backsubstitution" dopo
calcolare $\underline{L}^{-1}$. Consideriamo $N=4$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{10} & 1 & 0 & 0 \\ l_{20} & l_{21} & 1 & 0 \\ l_{30} & l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \Rightarrow \begin{cases} X_0 = x_0 \\ l_{10}X_0 + X_1 = x_1 \\ l_{20}X_0 + l_{21}X_1 + X_2 = x_2 \\ l_{30}X_0 + l_{31}X_1 + l_{32}X_2 + X_3 = x_3 \end{cases}$$

In generale:

$$\begin{cases} X_0 = x_0 \\ X_i = x_i - \sum_{\ell=0}^{i-1} l_{i\ell} X_\ell \quad i=1, \dots, N-1 \end{cases}$$

Questo porta alla struttura TRIANGOLARE INVERSA
mostata in figura



FATTORIZZAZIONE SPETTRALE
 Un'altra caratteristica di R_x molto importante deriva dalle soluzioni del problema agli autovalori per R_x . Mettendo insieme autovalori e autovettori otteniamo la seguente scomposizione

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

$$\left(\begin{array}{|c|} \hline \text{N autovettori} \\ \hline \end{array} \right) \left(\begin{array}{|c|} \hline \text{autovaleori} \\ \hline \end{array} \right) \left(\begin{array}{|c|} \hline \text{N autovettori} \\ \hline \end{array} \right) \left\{ \begin{array}{l} \text{N autovettori} \\ \text{coniugati} \end{array} \right.$$

autovaleori
nelle diagonali

La ortogonalità degli autovettori implica che $\underline{Q}^H \underline{Q} = \underline{I}_N$
 quindi $\underline{C} = \underline{Q}^{-1} = \underline{Q}^H$ ($\underline{Q}$ è matrice unitaria)

$$\underline{x} \rightarrow \boxed{\underline{Q}^H} \rightarrow \underline{X} \quad \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix} = \underline{Q}^H \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix}$$

Le componenti di $\underline{X}$ sono le proiezioni di $\underline{x}$ sugli autovettori di $\underline{R}_x$. La matrice di autocorrelazione di $\underline{X}$ è $E[\underline{X} \underline{X}^H] = \underline{Q}^H E[\underline{x} \underline{x}^H] \underline{Q} = \underline{Q}^H \underline{R}_x \underline{Q} = \underline{Q}^H \underline{Q} \underline{\Lambda} \underline{Q} \underline{Q}^H = \underline{\Lambda}$
 le varianze delle uscite sono proprio gli autovalori di $\underline{R}_x$.

Se scegliamo invece trasformata $\underline{C} = \underline{\Lambda}^{-1/2} \underline{Q}^H$,

otteniamo

$$E[\underline{X} \underline{X}^H] = \underline{\Lambda}^{-1/2} \underline{Q}^H E[\underline{x} \underline{x}^H] \underline{Q} \underline{\Lambda}^{-1/2} = \underline{\Lambda}^{-1/2} \underline{Q}^H \underline{Q} \underline{\Lambda} \underline{Q} \underline{Q}^H \underline{\Lambda}^{-1/2} = \underline{I}_N$$

ovvero che le componenti di $\underline{X}$ oltre a essere scalari, hanno anche varianza unitaria.

Nella trasformazione $\underline{C} = \underline{Q}^H$ ~~otteniamo~~ la proprietà di TDO
conservazione dell'energia, ovvero

$$\sum_{i=0}^{N-1} E[X_i^2] \triangleq \text{tr}[\underline{R}_x] = \sum_{i=0}^N E[X_i^2] = \text{tr}[E[\underline{X}\underline{X}^H]] = \sum_{i=0}^{N-1} \lambda_i$$

Lo prova e' immediato osservando che

$$\text{tr}[\underline{R}_x] = \text{tr}[\underline{Q}\underline{\Lambda}\underline{Q}^H]$$

Dalla proprietà $\text{tr}[AB] = \text{tr}[BA]$, abbiamo

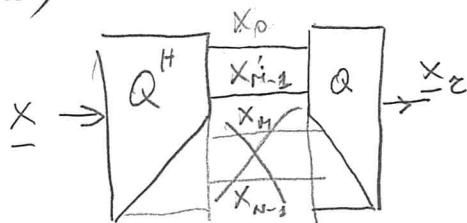
$$\text{tr}[\underline{Q}\underline{\Lambda}\underline{Q}^H] = \text{tr}[\underline{Q}^H \underline{Q} \underline{\Lambda}] = \text{tr}[\underline{\Lambda}] = \sum_{i=0}^{N-1} \lambda_i$$

Quindi la decomposizione spettrale ci fornisce una
 nuova rappresentazione a componenti scalari con
 una nuova distribuzione dell'energia. Questo è
 particolarmente interessante perché ci consente di
 volutare in quali direzioni, (ovvero nelle direzioni di quali autovettori
 si concentra la maggiore energia. La rappresentazione
 è utile anche per notare matrici $\underline{R}_x$ a rang. non
 pieno in cui alcuni degli autovalori sono nulli.

In generale la trasformazione è conservativa, poiché
 $\underline{x}$ è ricostruibile da $\underline{x}_z$ mediante una trasformazione
 inversa

$$\underline{x} \xrightarrow{\underline{Q}^H} \underline{x}_z \xrightarrow{\underline{Q}} \underline{x}_z = \underline{x}$$

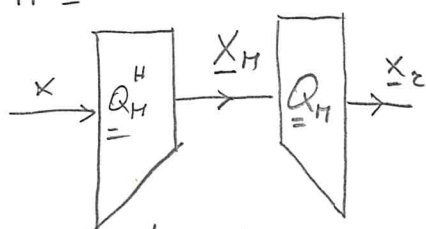
Ma vogliamo porci il problema di studiare cosa
 succede se solo $M < N$ componenti di $\underline{x}$ vengono
 conservate, ovvero studiare la ricostruzione di $\underline{x}$
 che ne consegue.



A questo scopo è utile partizionare la decomposizione
matriciale in due insieme:

$$R_X = \underbrace{Q_N}_{\substack{N \\ \parallel \dots \parallel}} \underbrace{\Lambda_N}_{\substack{N \\ \diagdown}} \underbrace{Q_N^H}_{\substack{N \\ \equiv}} = \underbrace{[Q_H \quad Q_{N-H}]}_{\substack{H \quad N-H}} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \underbrace{\begin{bmatrix} Q_H^H \\ Q_{N-H}^H \end{bmatrix}}_{\substack{H \\ \equiv \\ N-H}} = \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \vdots \\ \text{---} \end{bmatrix} \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \vdots \\ \text{---} \end{bmatrix}$$

La trasformazione basata solo sulle prime H componenti
è $\underline{X}_H = Q_H^H \underline{X}$ e la ricostruzione $\underline{X}_r = Q_H \underline{X}_H = Q_H Q_H^H \underline{X}$



Ora la domanda è: cosa si è perso nella trasforma-
zione? Ovvero di quanto differisce $\underline{X}_r$ da $\underline{X}$?

Definendo un vettore di errore $\underline{\varepsilon} = \underline{X} - \underline{X}_r$,

possiamo valutare la media $E[\underline{\varepsilon}] = E[\underline{X}] - E[\underline{X}_r] = 0$,

il valore quadratico medio $E[\underline{\varepsilon}^H \underline{\varepsilon}]$ e la

autocovarianza $E[\underline{\varepsilon} \underline{\varepsilon}^H]$. Ovviamente

il valore quadratico medio è la traccia di $E[\underline{\varepsilon} \underline{\varepsilon}^H]$,

per tanto valutiamo $E[\underline{\varepsilon} \underline{\varepsilon}^H]$.

$$\underline{R}_\varepsilon = E[\underline{\varepsilon} \underline{\varepsilon}^H] = E[(\underline{X} - \underline{X}_r)(\underline{X} - \underline{X}_r)^H]$$

$$= E[\underline{X} \underline{X}^H] + E[\underline{X}_r \underline{X}_r^H] - E[\underline{X}_r \underline{X}^H] - E[\underline{X} \underline{X}_r^H]$$

$$= \underline{R}_X + E[Q_H Q_H^H \underline{X} \underline{X}^H Q_H Q_H^H] - E[Q_H Q_H^H \underline{X} \underline{X}^H] - E[\underline{X} \underline{X}^H Q_H Q_H^H]$$

$$= R_x + Q_H Q_H^H R_x Q_H Q_H^H - Q_H Q_H^H R_x - R_x Q_H Q_H^H$$

$$= R_x + Q_H Q_H^H \begin{bmatrix} Q_H & Q_{N-H} \end{bmatrix} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{bmatrix} Q_H^H \\ Q_{N-H}^H \end{bmatrix} Q_H Q_H^H$$

$$- Q_H Q_H^H \begin{bmatrix} Q_H & Q_{N-H} \end{bmatrix} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{bmatrix} Q_H^H \\ Q_{N-H}^H \end{bmatrix}$$

$$- \begin{bmatrix} Q_H & Q_{N-H} \end{bmatrix} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{bmatrix} Q_H^H \\ Q_{N-H}^H \end{bmatrix} Q_H Q_H^H$$

$$= R_x + Q_H \begin{bmatrix} I_H & 0_{N-H} \end{bmatrix} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{bmatrix} I_H \\ 0_{N-H}^T \end{bmatrix} Q_H^H$$

$$- Q_H \begin{bmatrix} I_H & 0_{N-H} \end{bmatrix} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{bmatrix} Q_H^H \\ Q_{N-H}^H \end{bmatrix}$$

$$- \begin{bmatrix} Q_H & Q_{N-H} \end{bmatrix} \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{bmatrix} I_H \\ 0_{N-H}^T \end{bmatrix} Q_H^H$$

$$= R_x + Q_H \Lambda_H Q_H^H - Q_H \begin{bmatrix} \Lambda_H & 0 \end{bmatrix} \begin{bmatrix} Q_H^H \\ Q_{N-H}^H \end{bmatrix} - \begin{bmatrix} Q_H & Q_{N-H} \end{bmatrix} \begin{bmatrix} \Lambda_H \\ 0 \end{bmatrix} Q_H^H$$

$$= R_x + Q_H \Lambda_H Q_H^H - Q_H \Lambda_H Q_H^H - Q_H \Lambda_H Q_H^H$$

$$R_E = R_X - Q_H \Lambda_H Q_H^H$$

$$\begin{bmatrix} N \\ \square \end{bmatrix} \begin{bmatrix} N \\ \square \end{bmatrix} - \begin{bmatrix} M \\ \square \end{bmatrix} \begin{bmatrix} M \\ \square \end{bmatrix} \begin{bmatrix} M \\ \square \end{bmatrix}$$

Riscrivendo $Q_H \Lambda_H Q_H^H = [Q_H Q_{N-H}] \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{pmatrix} Q_H^H \\ Q_{N-H}^H \end{pmatrix}$

otteniamo

$$R_E = [Q_H Q_{N-H}] \begin{pmatrix} \Lambda_H & 0 \\ 0 & \Lambda_{N-H} \end{pmatrix} \begin{pmatrix} Q_H^H \\ Q_{N-H}^H \end{pmatrix} = Q_{N-H} \Lambda_{N-H} Q_{N-H}^H$$

$$\begin{matrix} N \\ \square \end{matrix} \begin{matrix} N-H \\ \square \end{matrix} \begin{matrix} N-H \\ \square \end{matrix}$$

Se $\Lambda_{N-H} \neq 0$ gli errori non sono correlati.

L'errore quadratico medio è

$$E[\varepsilon^H \varepsilon] = \text{tr}(R_E) = \text{tr}(Q_{N-H} \Lambda_{N-H} Q_{N-H}^H) = \text{tr}(\overbrace{Q_{N-H}^H Q_{N-H}}^{I_{N-H}} \Lambda_{N-H})$$

$$= \sum_{i=M}^{N-1} \lambda_i$$

L'errore quadratico medio è pari alla somma degli autovalori delle componenti trascurate.

Questo risultato suggerisce di ordinare gli autovalori in ordine decrescente

$$\lambda_0 > \lambda_1 > \dots > \lambda_{M-1} > \lambda_M > \dots > \lambda_{N-1}$$

Pertanto, conservando le prime M componenti si ottiene la migliore ricostruzione, ovvero la migliore compressione dell'energia.

Questo trasformare è anche noto come

o TRASFORMATA DI KARBUNEN-LOEVE DISCRETA

o KLT, o DKLT. Le componenti preservate sono anche dette COMPONENTI PRINCIPALI o Canalini (PRINCIPAL COMPONENTS)

e' detta ANALISI ALLE COMPONENTI PRINCIPALI (PCA - PRINCIPAL COMPONENT ANALYSIS)

Una buona collezione dell'analisi appena presentata, ^{7D15}
 che se R_x è a rango degenere, diciamo r ,
 solo r autovalori saranno diversi da zero.
 Pertanto prendendo solo le prime r componenti
 si ottiene una ricostruzione perfetta.

Nelle applicazioni, tipiche distribuzioni degli
 autovalori sono concentrate a ordini bassi
 per cui è possibile "comprimere" il vettore
 $\underline{x}$ in un numero limitato di componenti
 conservando gran parte dell'energia.

La DKL è molto utilizzata nell'elaborazione numerica
 dei segnali e in particolare nella elaborazione delle
 immagini.

A conclusione di questo discorso ricordiamo che
 le componenti principali, ovvero gli autovettori
 corrispondenti agli autovalori più grandi, sono
 ottenibili anche come soluzione del seguente
 problema:

Dato un vettore aleatorio $\underline{x}$ con matrice di autocorrelazione
 $R_x = E[\underline{x}\underline{x}^H]$, e dato un insieme di indici $k=0, \dots, L-1$,
 si trovi la trasformazione ortogonale decodificante A
 $\underline{X} = \underline{A} \underline{x}$, $\underline{A} \underline{A}^H = \underline{I}_L$ tale che l'energia
 $L \left\| \underline{X} \right\|_N^2 = \left\| \underline{x} \right\|_N^2$ $\sum_{k=0}^{L-1} |X[k]|^2$ sia massima.

$$\underline{A} = \begin{bmatrix} \underline{a}_0^H \\ \underline{a}_1^H \\ \vdots \\ \underline{a}_{L-1}^H \end{bmatrix}$$

Vediamo prima per $L=1$

TD14

$$X[0] = \underline{a}_0^H \underline{x}$$

Il problema è

$$\begin{cases} \underline{a}_0 = \underset{\underline{a}_0}{\operatorname{argmax}} E[|X[0]|^2] \\ \underline{a}_0^H \underline{a}_0 = 1 \end{cases}$$

$$E[|X[0]|^2] = \underline{a}_0^H E[\underline{x} \underline{x}^H] \underline{a}_0 = \underline{a}_0^H \underline{R}_x \underline{a}_0$$

Utilizzando la funzione di Lagrange

$$F(\underline{a}_0, \lambda) = \underline{a}_0^H \underline{R}_x \underline{a}_0 + \lambda (1 - \underline{a}_0^H \underline{a}_0)$$

mettendo a zero il gradiente

$$\nabla_{\underline{a}_0} F = 2 \underline{R}_x \underline{a}_0 - 2 \lambda \underline{a}_0 = \underline{0}$$

obteniamo

$$\underline{R}_x \underline{a}_0 = \lambda_0 \underline{a}_0 \quad \text{problema dell'autovettore massimo}$$

$$\lambda_0 = \frac{\underline{a}_0^H \underline{R}_x \underline{a}_0}{\underline{a}_0^H \underline{a}_0} \quad \text{equivalente a massimizzare questa quantità detta anche coefficiente di Rayleigh.}$$

$X[0]$ è la componente principale

Le altre componenti sono ottenute considerando lo stesso problema con vincoli di ortogonalità. Per esempio, per la seconda componente principale $\underline{a}_1$

$$\begin{cases} \underline{a}_1 = \underset{\underline{a}_1}{\operatorname{argmax}} E[|X[1]|^2] \\ \underline{a}_1^H \underline{a}_1 = 1 \\ \underline{a}_1^H \underline{a}_0 = 0 \end{cases}$$

$$\text{ovvero massimizzare } \lambda_1 = \frac{\underline{a}_1^H \underline{R}_x \underline{a}_1}{\underline{a}_1^H \underline{a}_1} \quad \text{con } \underline{a}_1^H \underline{a}_0 = 0$$

eccetera...