TRASFORMATE DISCRETE APPLICATE ALLE IMMAGINI

Lezioni del corso di
TRASMISSIONE ED ELABORAZIONE NUMERICA DEI SEGNALI

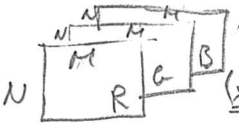
A.A. 2015-16

Prof. FRANCESCO A.N. PALMIERI

Nella elaborazione delle immagini molte delle tecniche introdotte per i segnali monodimensionali vengono estese con similarità formali. Ricordiamo che una immagine è una matrice di $N \times M$ pixel

$$\begin{array}{cc} & m \rightarrow \\ n \downarrow & \underline{\underline{X}} = \begin{bmatrix} x[0,0] & x[0,1] & \dots & x[0,M-1] \\ x[1,0] & x[1,1] & \dots & x[1,M-1] \\ \vdots & & & \\ x[N-1,0] & x[N-1,1] & \dots & x[N-1,M-1] \end{bmatrix} \end{array}$$

I valori dei pixel $\{x[n,m]\}$ possono essere reali o complessi. Si tratta ovviamente di una immagine di intensità (o B/W)

Le immagini a colori vengono tipicamente rappresentate con un insieme di 3 immagini 

, tipicamente RGB $(\underline{\underline{X}}_R, \underline{\underline{X}}_G, \underline{\underline{X}}_B)$

(esistono anche altre rappresentazioni quali per esempio LIQ - Luminanza Crominanza, o altri colori base)

Tralasceremo questo aspetto e considereremo in seguito solo immagini singole

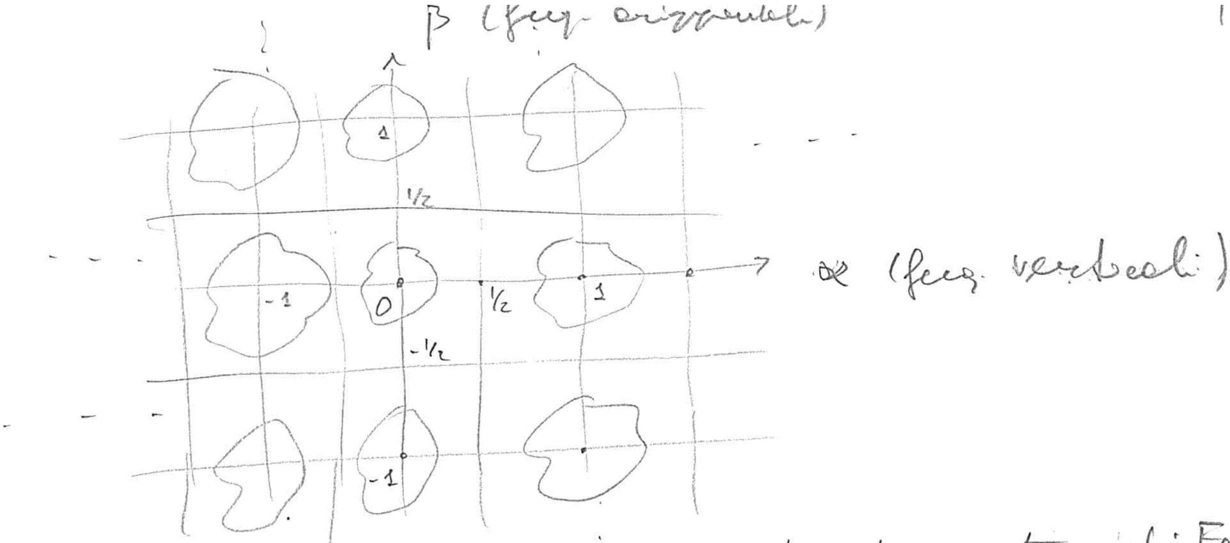
Se una immagine avesse estensione infinita

$$\{\underline{\underline{X}}[n,m]\}_{n=-\infty,+\infty}^{m=-\infty,+\infty}$$

potremo considerare la trasformata di Fourier 2D-DTFT

$$X(\alpha,\beta) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} x[n,m]\, e^{-j2\pi(\alpha n+\beta m)}$$

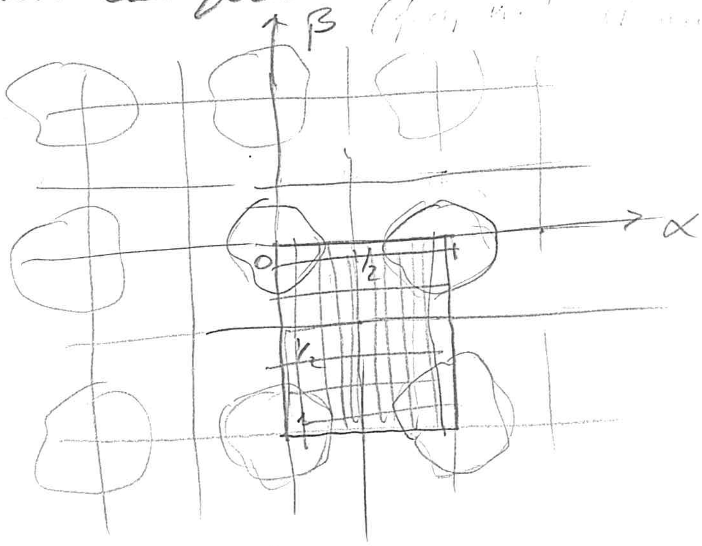
con α e β reali rappresentanti rispettivamente la frequenza orizzontale e la frequenza verticale normalizzata. La funzione $X(\alpha,\beta)$ è una funzione bi-dimensionale complessa definita nel campo reale $(\alpha,\beta) \in \mathbb{R}^2$ periodica sia in α che in β con periodo 1 come schematicamente mostrato in figura



La 2D-DTFT è ottenibile come la trasformata di Fourier bidimensionale continua di una funzione 2D continua campionata (campionamento 2D discusso in un'altra lezione)
Ora se consideriamo una sequenza 2D (immagine finita), la 2D-DTFT diventa

$$X(\alpha, \beta) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[n,m]\, e^{-j 2\pi (\alpha n + \beta m)}$$

che ha ovviamente le stesse caratteristiche mostrate in figura con la differenza che contiene solo $N \times M$ gradi di libertà. Pertanto essa può essere campionata con $N \times M$ campioni in un periodo (2D) che ne preservino l'invertibilità



$$\underline{X}[k,l] = X(\alpha,\beta)\Big|_{\substack{\alpha = \frac{k}{N} \quad k=0,\dots,N-1 \\ \beta = \frac{l}{M} \quad l=0,\dots,M-1}}$$

Questo porta alla definizione

$$\boxed{X[k,\ell] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[n,m]\, e^{-j\frac{2\pi}{N}nk}\, e^{-j\frac{2\pi}{M}m\ell} \qquad \begin{array}{l} k=0,\dots,N-1 \\ \ell=0,\dots,M-1 \end{array}}$$

2D-DFT — TRASFORMATA DISCRETA DI FOURIER BI-DIMENSIONALE

I valori della trasformata possono essere anch'essi organizzati in una matrice $N\times M$

$$\underline{X} = \begin{array}{c|c} & \ell \rightarrow \\ \hline k \downarrow & \begin{bmatrix} X[0,0] & X[0,1] & \dots & X[0,M-1] \\ X[1,0] & X[1,1] & \dots & X[1,M-1] \\ \vdots & & & \\ X[N-1,0] & X[N-1,1] & \dots & X[N-1,M-1] \end{bmatrix} \end{array}$$

Si noti dalla figura la posizione in frequenza della finestra su cui si effettua il campionamento: dello spettro della 2D-DTFT è nei quattro angoli (*)

In una immagine reale $X(\alpha,\beta)$ è Hermitiana, rispetto a entrambi le frequenze (immediato dalla definizione) $X(\alpha,\beta) = X^*(-\alpha,-\beta)$, quindi tutta l'informazione è già contenuta nel primo quarto della matrice

$$\underline{X} = \begin{array}{c|c} & M/2 \\ \hline \frac{N}{2} & \begin{bmatrix} X[0,0] & X[0,1] & \dots & X[0,\frac{M}{2}-1] \\ X[1,0] & X[1,1] & \dots & X[1,\frac{M}{2}-1] \\ \vdots & & & \\ X[\frac{N}{2}-1,0] & X[\frac{N}{2}-1,1] & \dots & X[\frac{N}{2}-1,\frac{M}{2}-1] \end{bmatrix} \end{array}$$

(*) Per ottenere la trasformata compresa da $-\frac{1}{2}$ a $\frac{1}{2}$ sia per α che per β, bisogna riorganizzare i quattro quarti di $\underline{X}$. Lo schema di campionamento adottato è quello tipico anche della DFT monodimensionale ed è quello che porta poi all'algoritmo FFT.

Separabilità della 2D-DFT

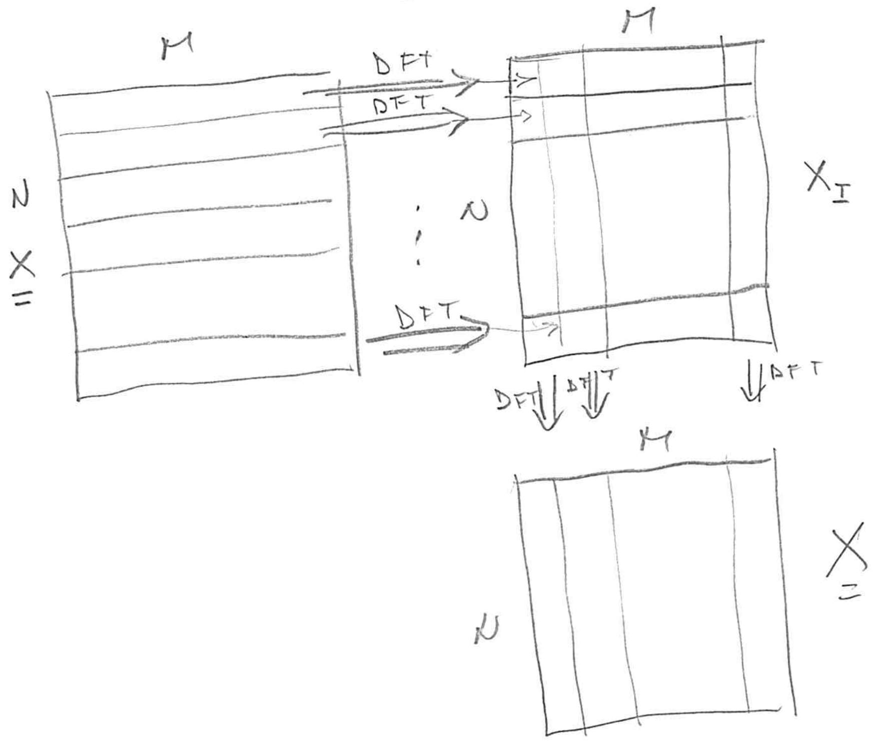
Notiamo che l'espressione della 2D-DFT può essere riorganizzata

$$X[k,l] = \sum_{n=0}^{N-1} e^{-j\frac{2\pi}{N}nk} \underbrace{\sum_{m=0}^{M-1} x[n,m]\, e^{-j\frac{2\pi}{M}ml}}_{\text{DFT della k-esima riga}}$$

$$X_I[n,l]$$

Quindi la 2D-DFT può essere calcolata in due passi:

1) Calcola la DFT (monodimensionale) di ogni riga e memorizzala in una matrice intermedia $\underline{X}_I$
2) Calcola la 2D-DFT calcolando le DFT (monodimensionali) sulle colonne di $\underline{X}_I$



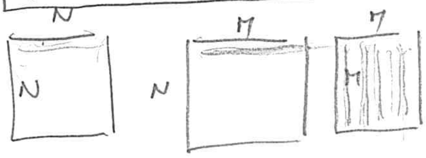
(si può ovviamente fare prima la trasformata per colonne e poi quella per righe)

Ovviamente ogni DFT monodimensionale può essere valutata usando un algoritmo veloce come la FFT.

La complessità computazionale è: $N \cdot M \log M$ per la prima fase e $M\,N \log N$ per la seconda. In totale la trasformata costa $O(MN \log N\cdot M)$ usando la FFT.

In forma matriciale l'operazione può essere riscritta come

$$\boxed{\underline{\underline{X}} = \underline{\underline{W}}_N \, \underline{\underline{x}} \, \underline{\underline{W}}_M}$$



dove $\underline{\underline{W}}_N = \left[e^{-j\frac{2\pi}{N}kn}\right]_{\substack{k=0..N-1\\ n=0..N-1}}$

$$\underline{\underline{W}}_M = \left[e^{-j\frac{2\pi}{M}lm}\right]_{\substack{l=0..M-1\\ m=0..M-1}}$$

Dalla conoscenza dell'invertibilità della DFT monodimensionale $\underline{\underline{W}}_N^{-1} = \frac{1}{N}\underline{\underline{W}}_N^*$, $\underline{\underline{W}}_M^{-1} = \frac{1}{M}\underline{\underline{W}}_M^*$, moltiplicando a sinistra per W_N^{-1} e a destra per W_M^{-1}, la 2D-DFT^{-1} può essere scritta come:

$$\boxed{\underline{\underline{x}} = \frac{1}{NM}\underline{\underline{W}}_N^* \, \underline{\underline{X}} \, \underline{\underline{W}}_M^*}$$

Ovviamente nella versione con le sommatorie si scrive

$$\boxed{x[n,m] = \frac{1}{NM}\sum_{l=0}^{M-1}\sum_{k=0}^{N-1} X[k,l]\, e^{j\frac{2\pi}{N}kn}\, e^{j\frac{2\pi}{M}lm} \qquad \begin{matrix} n=0..N-1\\ m=0,..M-1\end{matrix}}$$

2D-DFT^{-1}

TRASFORMATE 2D SEPARABILI

La semplice procedura per il calcolo della 2D-DFT deriva dalla separabilità del Kernel esponenziale

$$\underbrace{e^{-j\frac{2\pi}{N}nk}}\;\underbrace{e^{-j\frac{2\pi}{M}ml}}$$

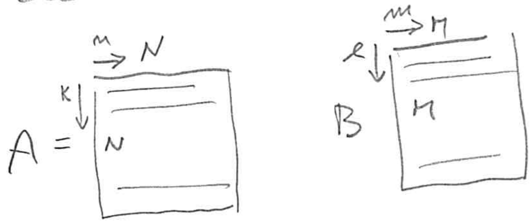
che consente di eseguire una trasformata 1D sulle righe e poi una sulle colonne. Questo porta anche alla espressione matriciale compatta.

Si possono pertanto definire delle trasformate 2D più generali separabili come dirette estensione di trasformate 1D.

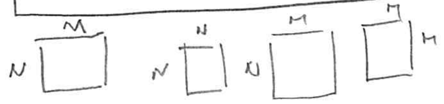
$$\boxed{\begin{gathered} X[k,l] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x[n,m]\, A[k,n]\, B[l,m] \\ k = 0,\dots N-1 \\ l = 0,\dots,M-1 \end{gathered}}$$

Questo equivale a operare la trasformazione con il Kernel $B[l,m]$ sulle righe e la trasformazione con il Kernel $A[k,n]$ sulle colonne della matrice risultante.

In forma matriciale assumendo che le righe delle due matrici contengano i vettori delle trasformate (nella DFT la matrice W_N è simmetrica quindi righe e colonne sono uguali)

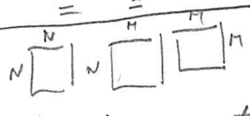


$$\boxed{\underline{\underline{X}} = \underline{\underline{A}}\;\underline{\underline{x}}\;\underline{\underline{B}}^T}$$



Per l'inversa

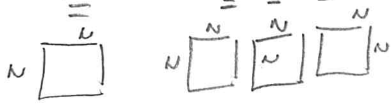
$$\boxed{\underline{\underline{x}} = \underline{\underline{A}}^{-1}\,\underline{\underline{X}}\,\underline{\underline{B}}^{-T}}$$



Quindi come trasformate 1D per A e per B possono essere utilizzate tutte quelle definite nelle precedenti lezioni.

Nel caso tipico in cui $A = B$ e $N = M$

$$\underline{\underline{X}} = \underline{\underline{A}}\,\underline{\underline{x}}\,\underline{\underline{A}}^T \quad ; \quad \underline{\underline{x}} = \underline{\underline{A}}^{-1}\underline{\underline{X}}\,\underline{\underline{A}}^{-T}$$

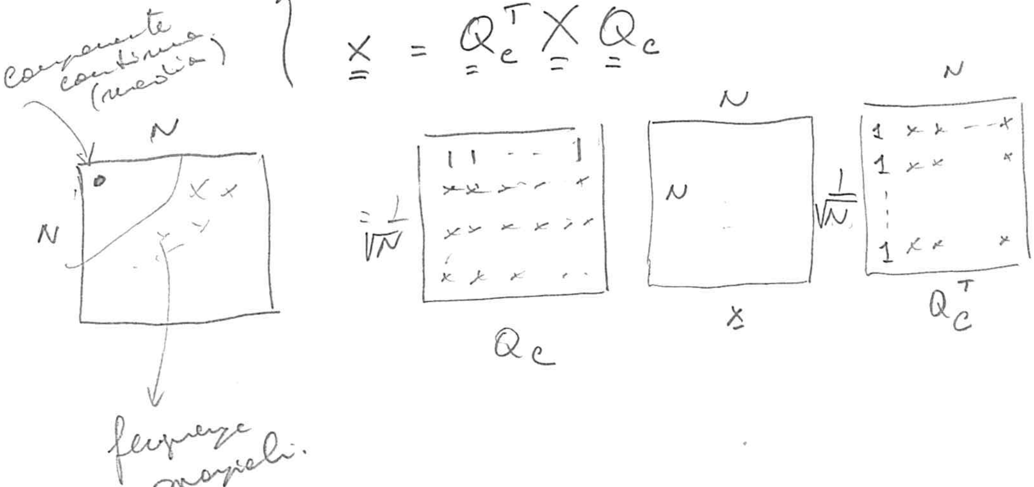


Per le trasformate ortonormali $A^{-1} = A^H$, quindi

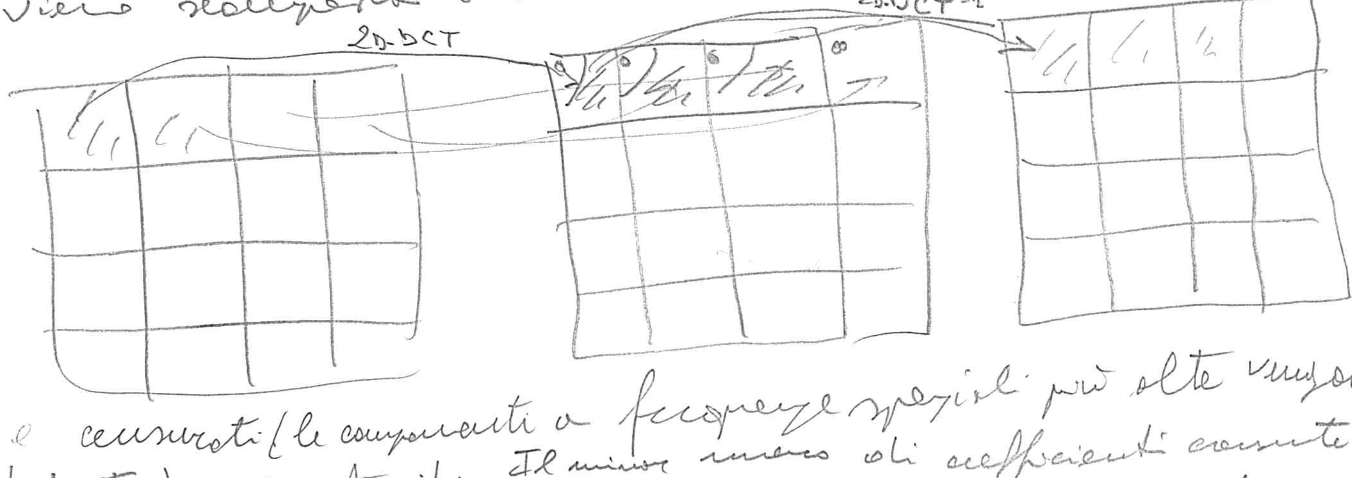
$$\underline{\underline{x}} = \underline{\underline{A}}^H \underline{\underline{X}}\,\underline{\underline{A}}^*$$

Esempio notevole è la estensione 2D della DCT in cui $\underline{\underline{A}} = \underline{\underline{Q}}_c$ con $\underline{\underline{Q}}_c^{-1} = \underline{\underline{Q}}_c^T$

$$\begin{cases} \underline{\underline{X}} = \underline{\underline{Q}}_c\,\underline{\underline{x}}\,\underline{\underline{Q}}_c^T \\ \underline{\underline{x}} = \underline{\underline{Q}}_c^T\,\underline{\underline{X}}\,\underline{\underline{Q}}_c \end{cases}$$



Il primo elemento in alto a sinistra è la media dell'immagine $\frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} x[n,m]$. Gli altri sono i contributi alle varie frequenze spaziali. Nella applicazione del JPEG per la compressione (con perdita) delle immagini, l'immagine viene scomposta in macroblocchi che vengono trasformati,



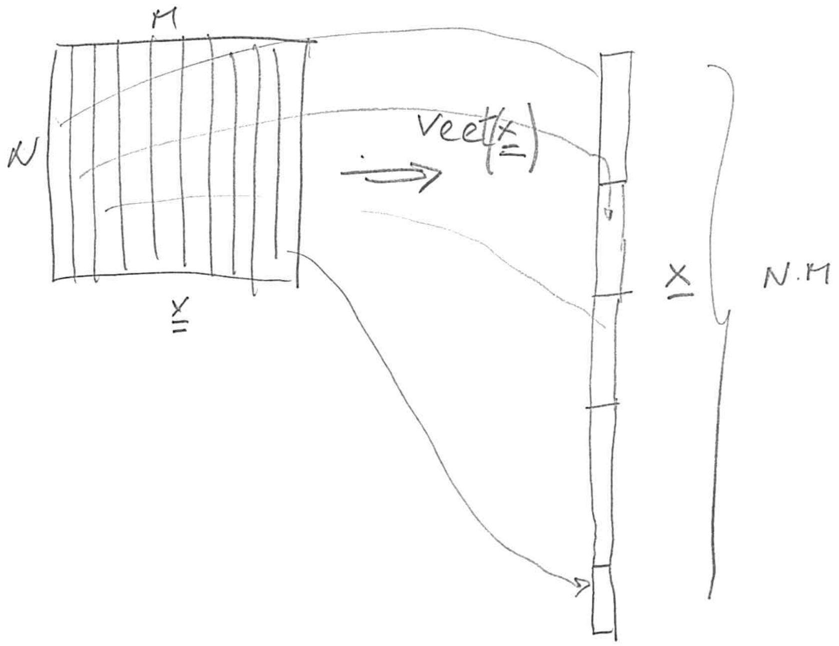
e censurati (le componenti a frequenze spaziali più alte vengono tagliate) e ricostruiti. Il minor numero di coefficienti consente una codifica più compatta con limitata perdita di informazione

TRASFORMATE NON SEPARABILI

In una trasformata non separabile

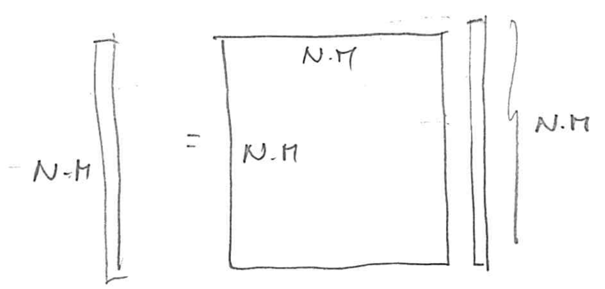
$$\boxed{X[k,\ell] = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} x[n,m]\ c[k,n,\ell,m] \qquad k=0,\dots N-1\ ,\ \ell=0,\dots,M-1}$$

il kernel $c[k,n,\ell,m]$ non è necessariamente fattorizzabile.
La versione matriciale della trasformata può essere formalizzata trasformando l'immagine in un vettore

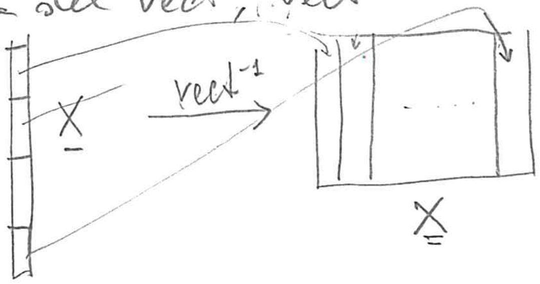


stacking all the columns of $\underline{\underline{x}}$ into a vector $\underline{x}$.

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



$\underline{X}$ può essere ridotta a matrice mediante l'operazione inversa del vect, $vect^{-1}$



Un esempio notevole è la KLT bi-dimensionale.
Si valuta la autocorrelazione di $\underline{x} = \text{vect}[\underline{\underline{X}}]$ e la si tratta
come se fosse un processo vettoriale di dimensioni N·M.
Il risultato poi si riduce a matrice mediante un vect^{-1}.

Nel caso di trasformate separabili

$$\underline{\underline{X}} = \underline{\underline{A}} \underline{\underline{x}} \underline{\underline{B}}^T$$

$$\underline{X} = \text{vect}\, \underline{\underline{X}} = (B \otimes A)\, \text{vect}\, \underline{\underline{x}}$$

$$\boxed{C = \underline{\underline{B}} \otimes \underline{\underline{A}}}$$

base property
$\text{vect}[ABC] = (C' \otimes A)\, \text{vect}(B)$
$\otimes$ Kronecker product.