Lezione 2B

Questa lezione spiega come sovradeterminare i sitemi di equazioni lineari.   
La formula qui sotto è la definizione formale, abbiamo un sistema lineare dove le equazioni sono in numero  
maggiore rispetto alle incognite, A è una matrice rettangolare e x è un vettore colonna. L'output è   
anch'esso un vettore colonna.  
Noi già sappiamo risolverlo, l'idea è trovare i coefficienti del polinomio che mi permettono di creare  
quella linea grafica che passa più vicino ai punti a disposizione.  
In pratica questo problema va sotto il nome di  
DATA FITTING.  
In ogni caso noi abbiamo A e b, dobbiamo  
trovare x che contiene i coefficienti del polinomio che fitta i dati.  
A è la matriche ce contiene tutte le coordinate X e b è il vettore che contiene le Y, l'obiettivo è trovare i coefficienti X

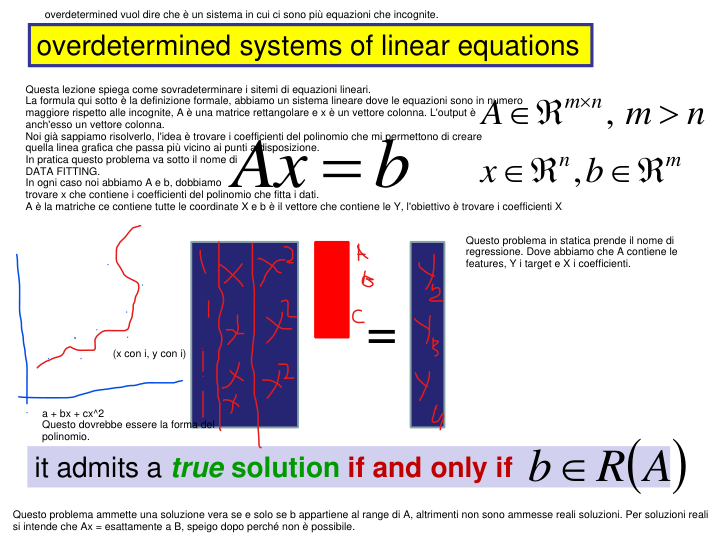
(x con i, y con i)

a + bx + cx^2  
Questo dovrebbe essere la forma del  
polinomio.

Questo problema in statica prende il nome di regressione. Dove abbiamo che A contiene le features, Y i target e X i coefficienti.

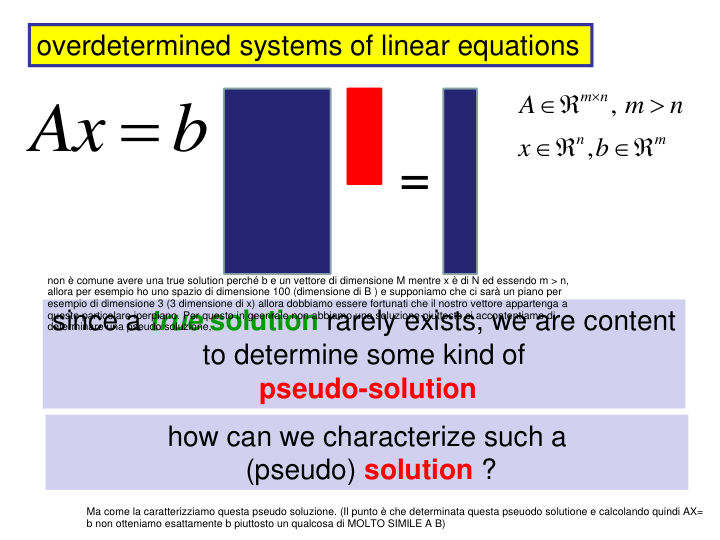
Questo problema ammette una soluzione vera se e solo se b appartiene al range di A, altrimenti non sono ammesse reali soluzioni. Per soluzioni reali si intende che Ax = esattamente a B, speigo dopo perché non è possibile.

overdetermined vuol dire che è un sistema in cui ci sono più equazioni che incognite.



non è comune avere una true solution perché b e un vettore di dimensione M mentre x è di N ed essendo m > n, allora per esempio ho uno spazio di dimensione 100 (dimensione di B ) e supponiamo che ci sarà un piano per esempio di dimensione 3 (3 dimensione di x) allora dobbiamo essere fortunati che il nostro vettore appartenga a questo particolare iperpiano. Per questo in geenrale non abbiamo una soluzione piuttosto ci accontentiamo di determinare una pseudo soluzione,

Ma come la caratterizziamo questa pseudo soluzione. (Il punto è che determinata questa pseuodo solutione e calcolando quindi AX= b non otteniamo esattamente b piuttosto un qualcosa di MOLTO SIMILE A B)

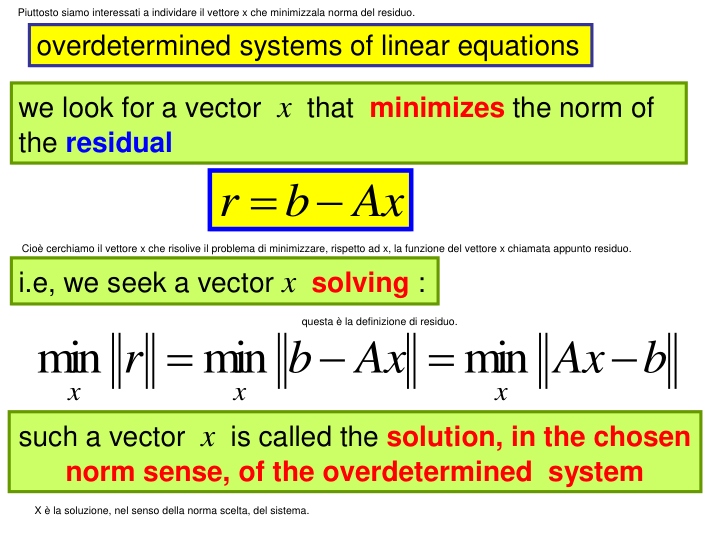


Piuttosto siamo interessati a individare il vettore x che minimizzala norma del residuo.

Cioè cerchiamo il vettore x che risolive il problema di minimizzare, rispetto ad x, la funzione del vettore x chiamata appunto residuo.

questa è la definizione di residuo.

X è la soluzione, nel senso della norma scelta, del sistema.

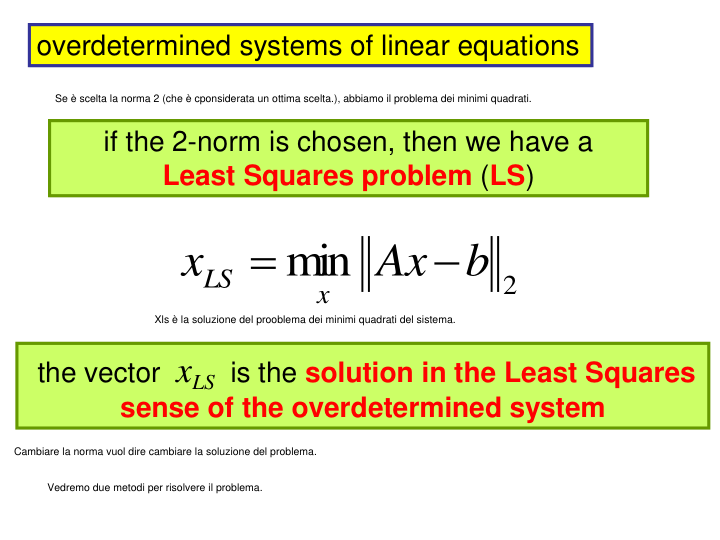


Se è scelta la norma 2 (che è cponsiderata un ottima scelta.), abbiamo il problema dei minimi quadrati.

Xls è la soluzione del prooblema dei minimi quadrati del sistema.

Vedremo due metodi per risolvere il problema.

Cambiare la norma vuol dire cambiare la soluzione del problema.



Questo dice è un esempio (strano non so perché)

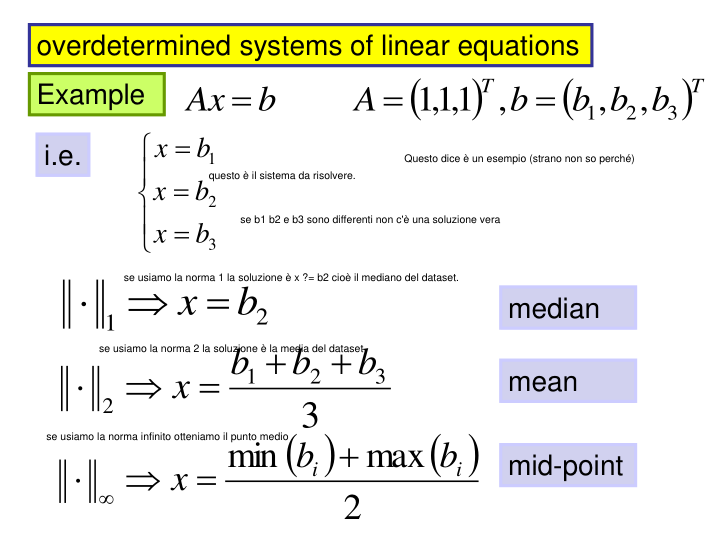
questo è il sistema da risolvere.

se b1 b2 e b3 sono differenti non c'è una soluzione vera

se usiamo la norma 1 la soluzione è x ?= b2 cioè il mediano del dataset.

se usiamo la norma 2 la soluzione è la media del dataset

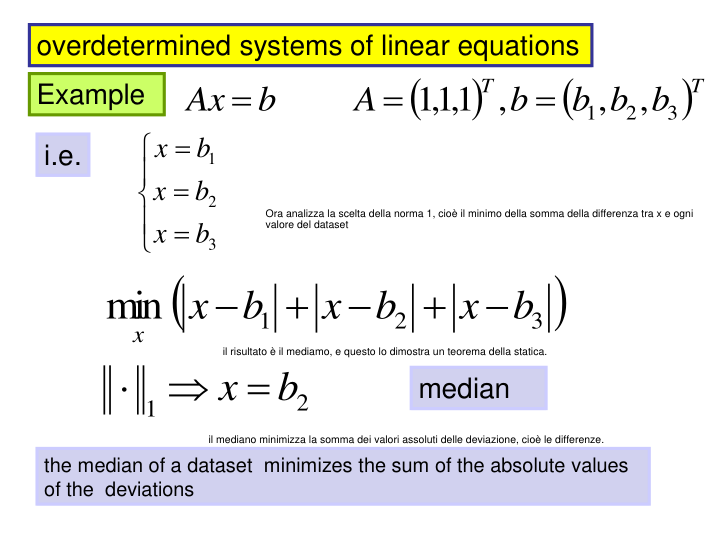
se usiamo la norma infinito otteniamo il punto medio



Ora analizza la scelta della norma 1, cioè il minimo della somma della differenza tra x e ogni valore del dataset

il risultato è il mediamo, e questo lo dimostra un teorema della statica.

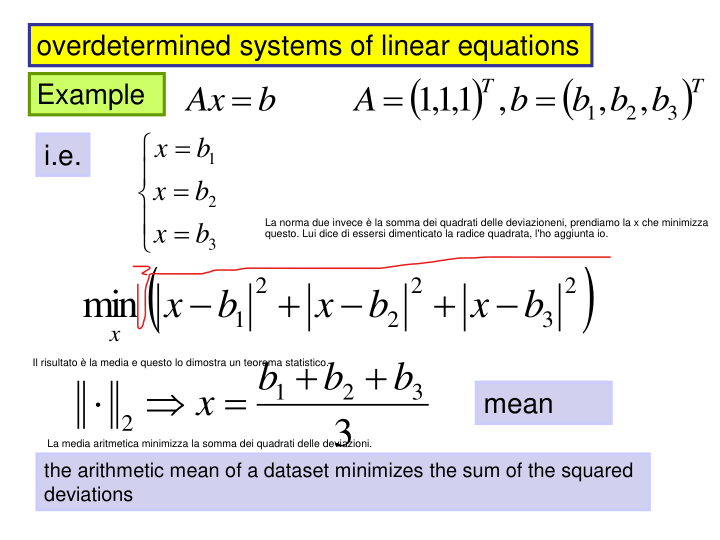
il mediano minimizza la somma dei valori assoluti delle deviazione, cioè le differenze.



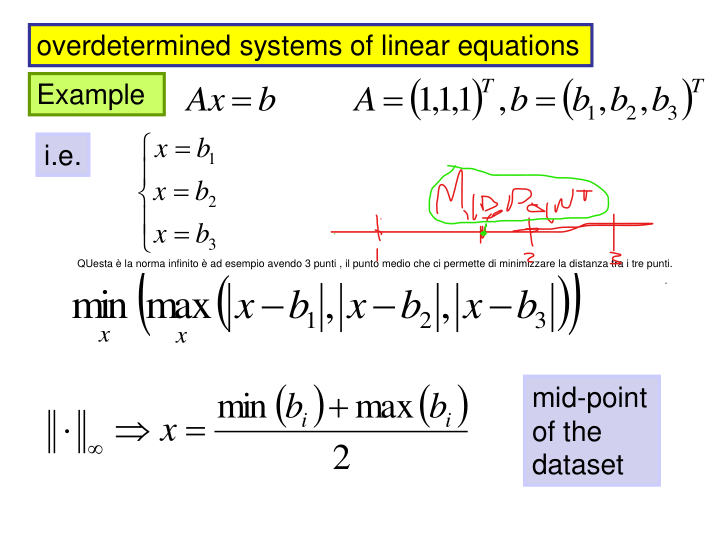
La norma due invece è la somma dei quadrati delle deviazioneni, prendiamo la x che minimizza questo. Lui dice di essersi dimenticato la radice quadrata, l'ho aggiunta io.

Il risultato è la media e questo lo dimostra un teorema statistico.

La media aritmetica minimizza la somma dei quadrati delle deviazioni.



QUesta è la norma infinito è ad esempio avendo 3 punti , il punto medio che ci permette di minimizzare la distanza tra i tre punti.



Vediamo alcune proprietà del problema dei minimi quadrati, non sappiamo quante soluzioni ha il sistema ne se le ha, questo dipende dalla metrica A.  
Noi chiamimao X come l'insieme che contiene le x che risolvono questo problema (minimi quadrati), si tratta di un insieme convesso.

Dato un insieme considerando due punti, se traccio la linea che collega due punti e la linea è nell'0insieme allora l'insieme è convesso altrimenti no

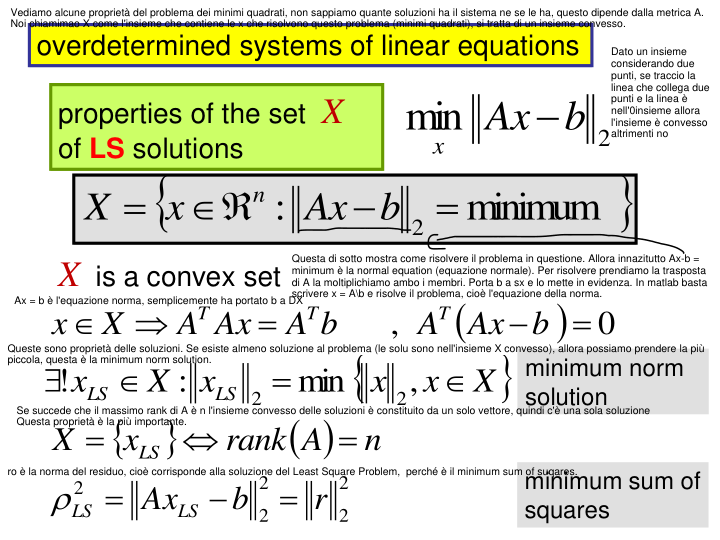
Questa di sotto mostra come risolvere il problema in questione. Allora innazitutto Ax-b = minimum è la normal equation (equazione normale). Per risolvere prendiamo la trasposta di A la moltiplichiamo ambo i membri. Porta b a sx e lo mette in evidenza. In matlab basta scrivere x = A\b e risolve il problema, cioè l'equazione della norma.

Ax = b è l'equazione norma, semplicemente ha portato b a DX

Queste sono proprietà delle soluzioni. Se esiste almeno soluzione al problema (le solu sono nell'insieme X convesso), allora possiamo prendere la più piccola, questa è la minimum norm solution.

Se succede che il massimo rank di A è n l'insieme convesso delle soluzioni è constituito da un solo vettore, quindi c'è una sola soluzione  
Questa proprietà è la più importante.

ro è la norma del residuo, cioè corrisponde alla soluzione del Least Square Problem, perché è il minimum sum of suqares.

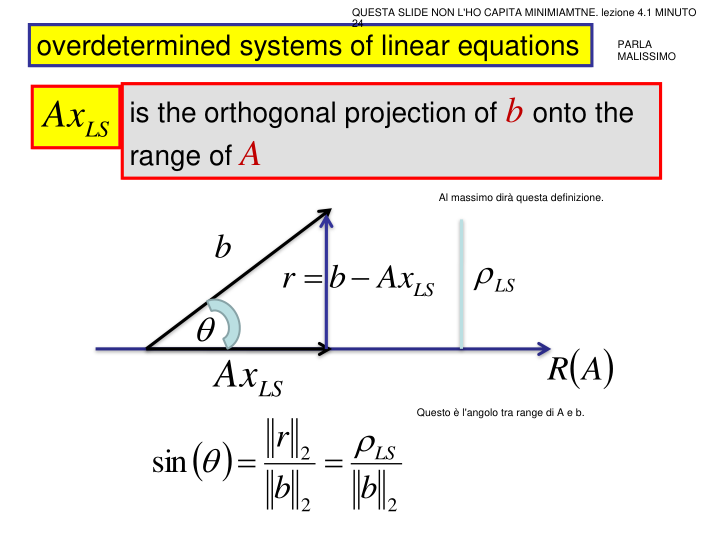


QUESTA SLIDE NON L'HO CAPITA MINIMIAMTNE. lezione 4.1 MINUTO 24

Al massimo dirà questa definizione.

Questo è l'angolo tra range di A e b.

PARLA MALISSIMO

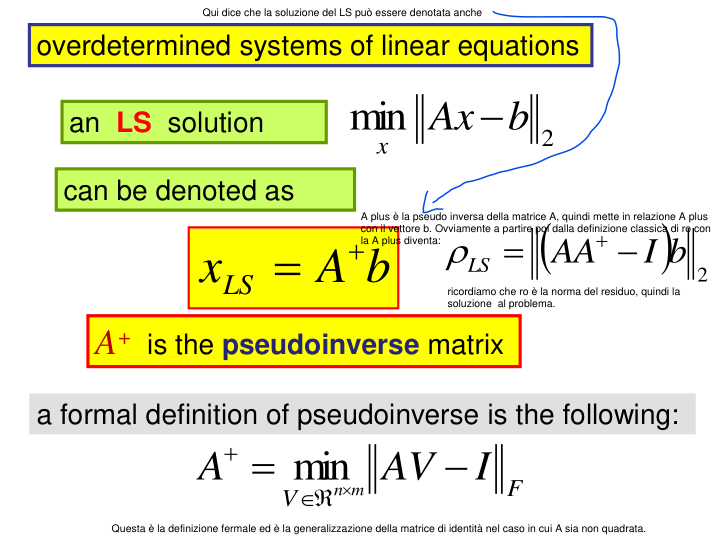


Qui dice che la soluzione del LS può essere denotata anche

A plus è la pseudo inversa della matrice A, quindi mette in relazione A plus con il vettore b. Ovviamente a partire poi dalla definizione classica di ro con la A plus diventa:

ricordiamo che ro è la norma del residuo, quindi la soluzione al problema.

Questa è la definizione fermale ed è la generalizzazione della matrice di identità nel caso in cui A sia non quadrata.

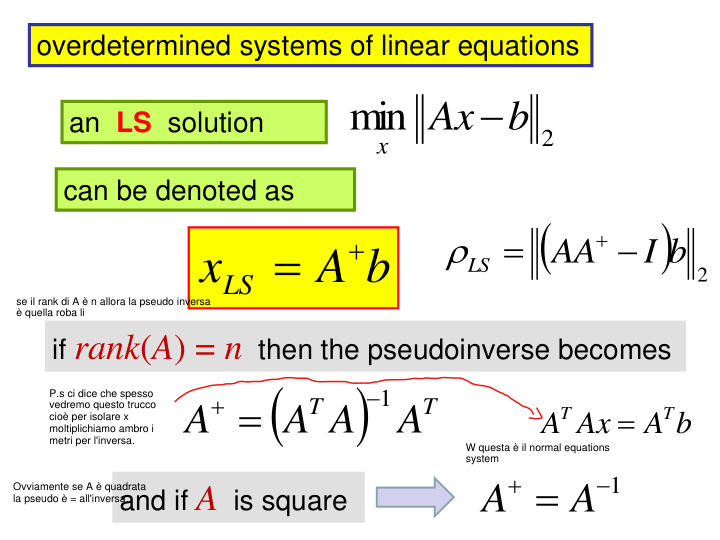


se il rank di A è n allora la pseudo inversa  
è quella roba li

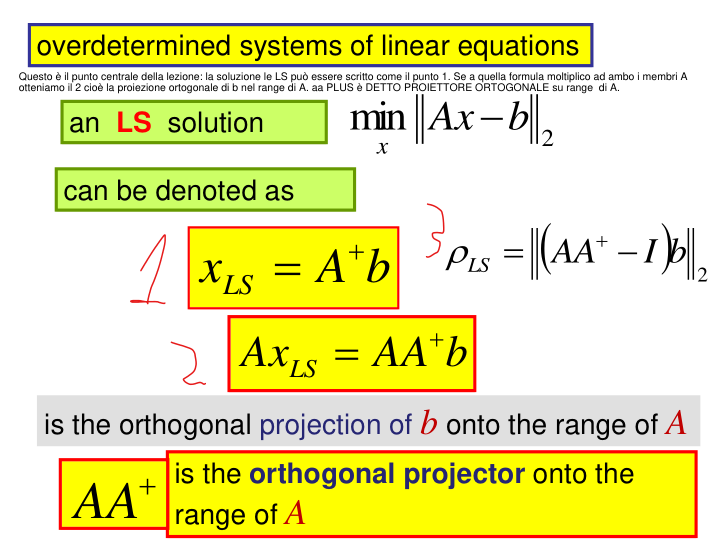
W questa è il normal equations system

P.s ci dice che spesso vedremo questo trucco cioè per isolare x moltiplichiamo ambro i metri per l'inversa.

Ovviamente se A è quadrata  
la pseudo è = all'inversa.



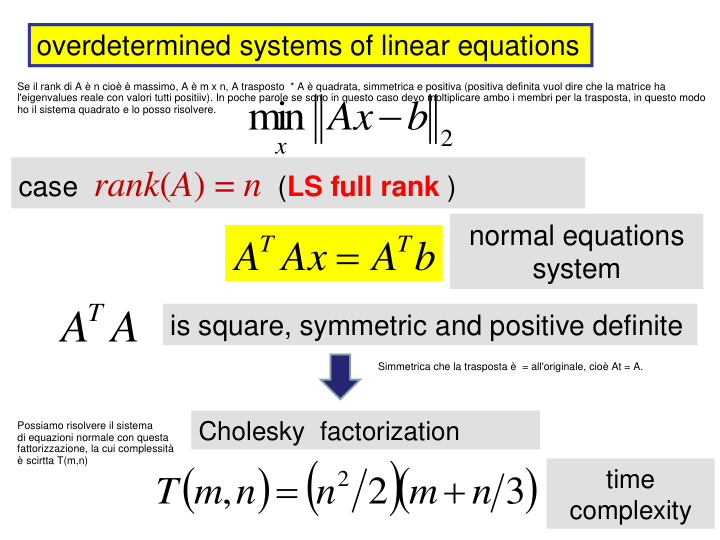
Questo è il punto centrale della lezione: la soluzione le LS può essere scritto come il punto 1. Se a quella formula moltiplico ad ambo i membri A otteniamo il 2 cioè la proiezione ortogonale di b nel range di A. aa PLUS è DETTO PROIETTORE ORTOGONALE su range di A.



Se il rank di A è n cioè è massimo, A è m x n, A trasposto \* A è quadrata, simmetrica e positiva (positiva definita vuol dire che la matrice ha l'eigenvalues reale con valori tutti positiiv). In poche parole se sono in questo caso devo moltiplicare ambo i membri per la trasposta, in questo modo ho il sistema quadrato e lo posso risolvere.

Simmetrica che la trasposta è = all'originale, cioè At = A.

Possiamo risolvere il sistema  
di equazioni normale con questa  
fattorizzazione, la cui complessità  
è scirtta T(m,n)



In questa slide che in questo caso analizzato, (lo stesso della slide precednete) c'è un indice di condizionamento molto grande. Perché l'indice di condizionamento della matrice (At \* Ax) è = al quadrato dell'indice di condizionamento di A. Se A è mal condizionato il problema radodpia questo mal condizionamtno.

l'indicie di cond di una matrice rett. è la norma 2 \* la norma 2 della matrice pseudo inversa

