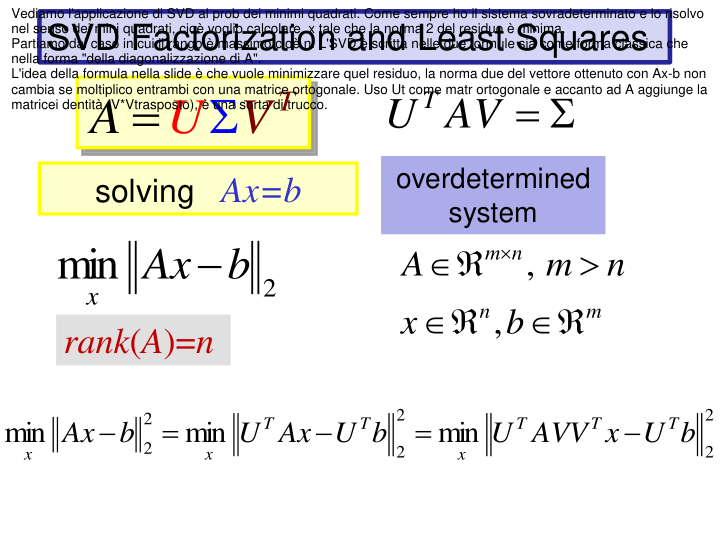
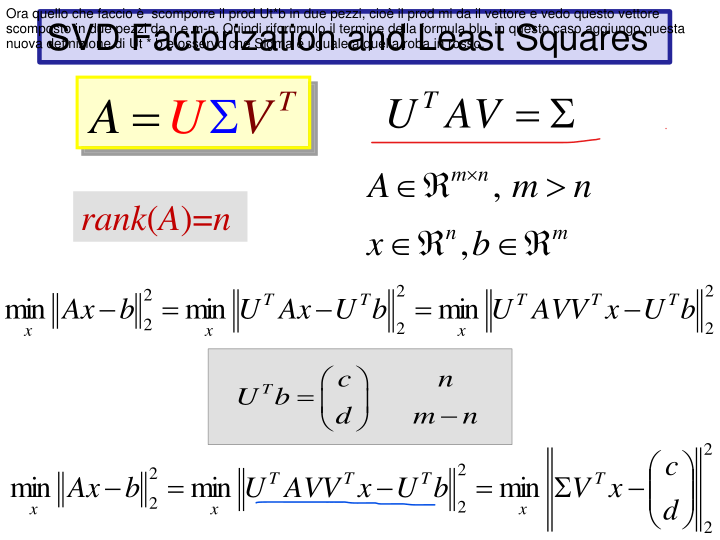
Lezione 6

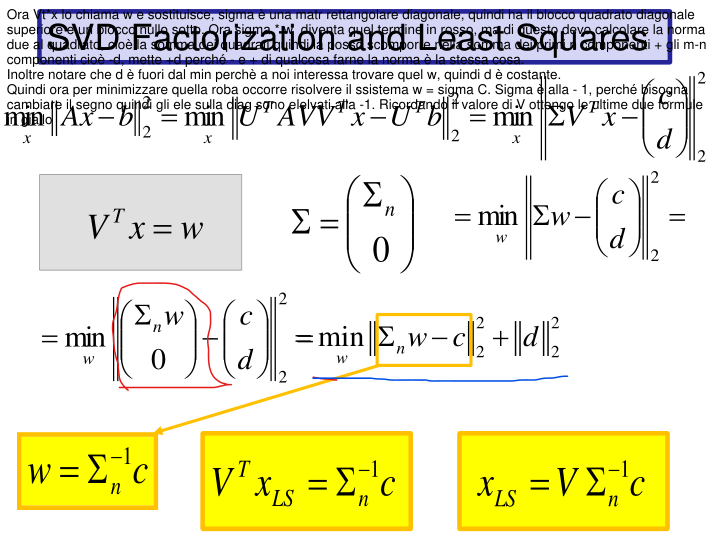
Vediamo l'applicazione di SVD al prob dei minimi quadrati. Come sempre ho il sistema sovradeterminato e lo risolvo nel senso dei mini quadrati, cioè voglio calcolare x tale che la norma 2 del residuo è minima.  
Partiamo dal caso in cui il rango è massimo cioè n. L'SVD è scritta nelle due formule sia come forma classica che nella forma "della diagonalizzazione di A".  
L'idea della formula nella slide è che vuole minimizzare quel residuo, la norma due del vettore ottenuto con Ax-b non cambia se moltiplico entrambi con una matrice ortogonale. Uso Ut come matr ortogonale e accanto ad A aggiunge la matricei dentità (V\*Vtrasposto), è una sorta di trucco.



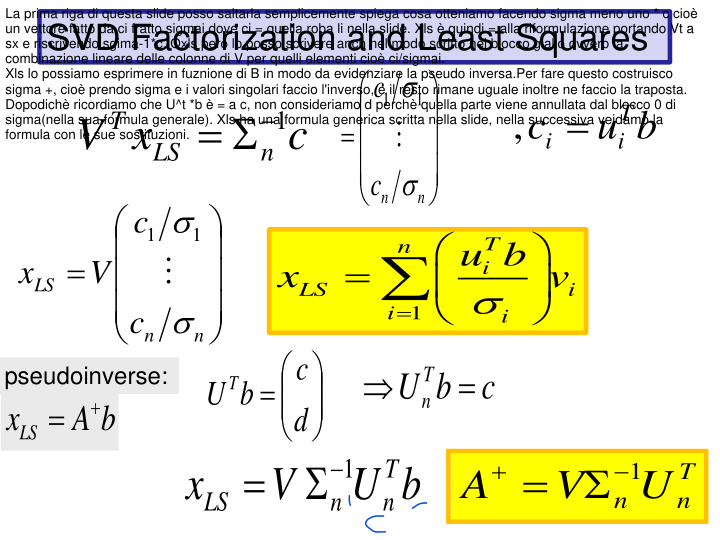
Ora quello che faccio è scomporre il prod Ut\*b in due pezzi, cioè il prod mi da il vettore e vedo questo vettore scomposto in due pezzi da n e m-n. Quindi riforumulo il termine della formula blu, in questo caso aggiungo questa nuova definizione di Ut \* b e osservo che Sigma é uguale a quella roba in rosso.



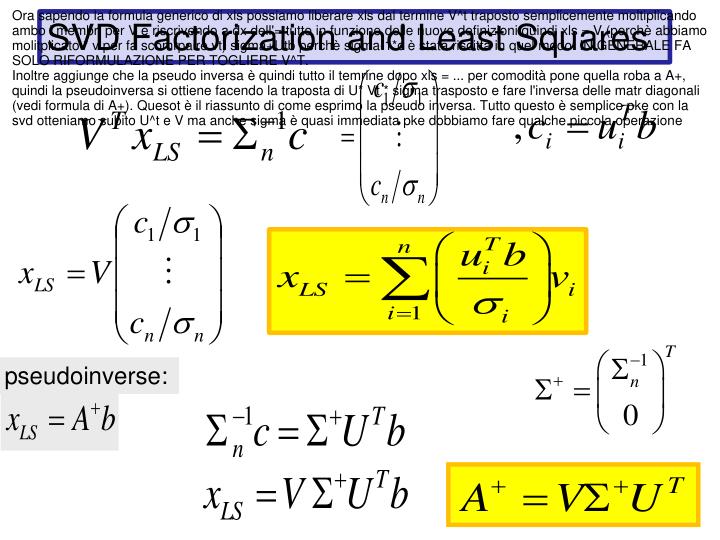
Ora Vt\*x lo chiama w e sostituisce, sigma è una matr rettangolare diagonale, quindi ha il blocco quadrato diagonale superiore e un blocco nullo sotto. Ora sigma \* w diventa quel termine in rosso, ma di questo devo calcolare la norma due al quadrato, cioè la somma dei quadrati quindi la posso scomporre nella somma dei primi n componenti + gli m-n componenti cioè -d, mette +d perché - e + di qualcosa farne la norma è la stessa cosa.  
Inoltre notare che d è fuori dal min perchè a noi interessa trovare quel w, quindi d è costante.  
Quindi ora per minimizzare quella roba occorre risolvere il ssistema w = sigma C. Sigma è alla - 1, perché bisogna cambiare il segno quindi gli ele sulla diag sono elelvati alla -1. Ricordando il valore di V ottengo le ultime due formule in giallo



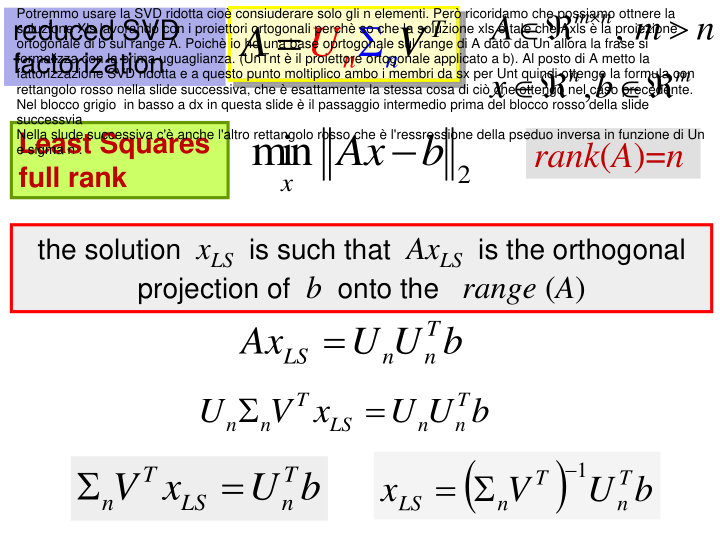
La prima riga di questa slide posso saltarla semplicemente spiega cosa otteniamo facendo sigma meno uno \* c cioè un vettore fatto da ci fratto sigmai dove ci = quella roba li nella slide. Xls è quindi = alla riformulazione portando Vt a sx e riscrivendo sgima-1\*c. Qxls però lo posso scrivere anch nel modo scritto nel blocco giallo ovvero la combinazione lineare delle colonne di V per quelli elementi cioè ci/sigmai.  
Xls lo possiamo esprimere in fuznione di B in modo da evidenziare la pseudo inversa.Per fare questo costruisco sigma +, cioè prendo sigma e i valori singolari faccio l'inverso, e il resto rimane uguale inoltre ne faccio la traposta. Dopodichè ricordiamo che U^t \*b è = a c, non consideriamo d perchè quella parte viene annullata dal blocco 0 di sigma(nella sua formula generale). Xls ha una formula generica scritta nella slide, nella successiva veidamo la formula con le sue sostituzioni.

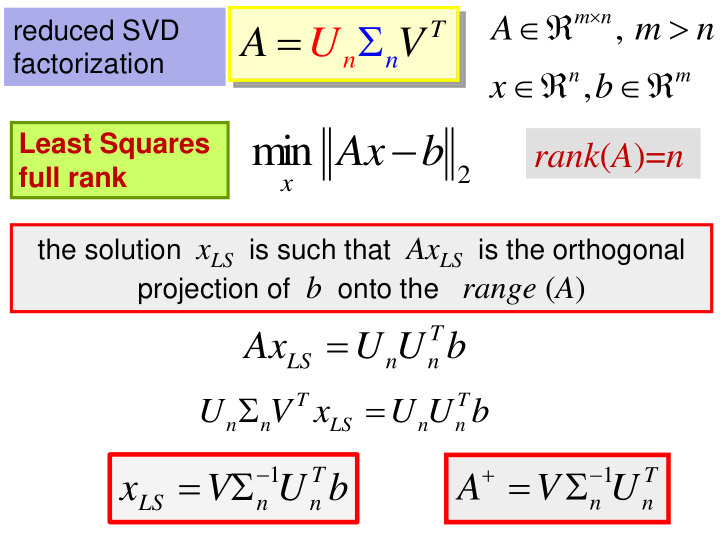


Ora sapendo la formula generico di xls possiamo liberare xls dal termine V^t traposto semplicemente moltiplicando ambo i membri per V e riscrivendo a dx dell'= tutto in funzione delle nuove definizioni quindi xls = V (perchè abbiamo molitplicato \* v per fa scomrpaire vt) sigma+Utb perchè sigma-1\*c è stata riscitta in quel modo. IN GENERALE FA SOLO RIFORMULAZIONE PER TOGLIERE V^T.  
Inoltre aggiunge che la pseudo inversa è quindi tutto il temrine dopo xls = ... per comodità pone quella roba a A+, quindi la pseudoinversa si ottiene facendo la traposta di U\* Vt \* sigma trasposto e fare l'inversa delle matr diagonali (vedi formula di A+). Quesot è il riassunto di come esprimo la pseudo inversa. Tutto questo è semplice pke con la svd otteniamo subito U^t e V ma anche sigma è quasi immediata pke dobbiamo fare qualche piccola operazione

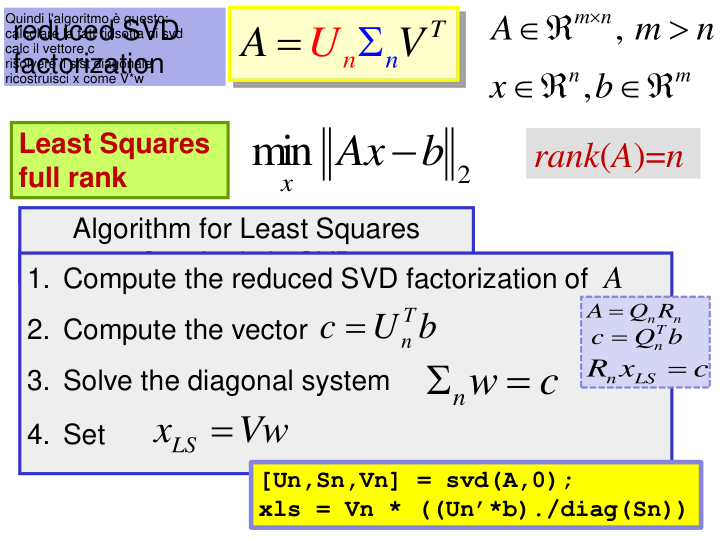


Potremmo usare la SVD ridotta cioè consiuderare solo gli n elementi. Però ricoridamo che possiamo ottnere la soluzione Xls lavorando con i proiettori ortogonali perchè so che la soluzione xls è tale che Axls è la proiezione ortogonale di b sul range A. Poichè io ho una base oprtogonale sul range di A dato da Un allora la frase si formalizza con la prima uguaglianza. (UnTnt è il proiettore ortogonale applicato a b). Al posto di A metto la fattorizzazione SVD ridotta e a questo punto moltiplico ambo i membri da sx per Unt quindi ottengo la formula con rettangolo rosso nella slide successiva, che è esattamente la stessa cosa di ciò che ottengo nel caso precedente. Nel blocco grigio in basso a dx in questa slide è il passaggio intermedio prima del blocco rosso della slide successvia  
Nella slude successiva c'è anche l'altro rettangolo rosso che è l'ressressione della pseduo inversa in funzione di Un e sigma n .

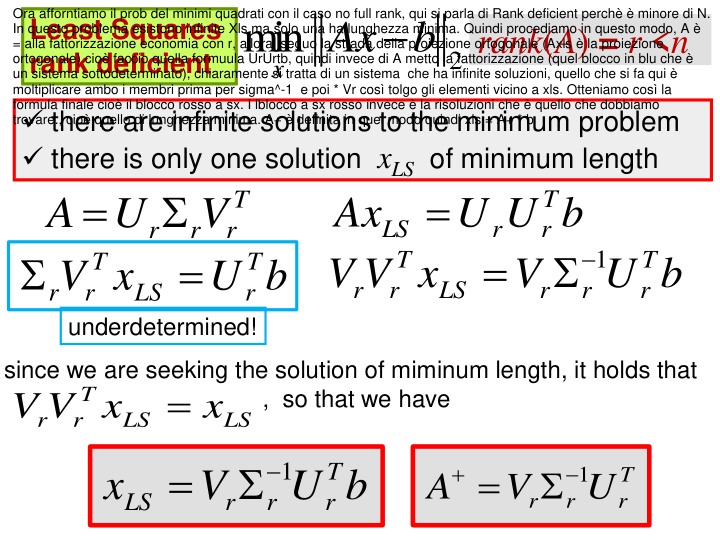




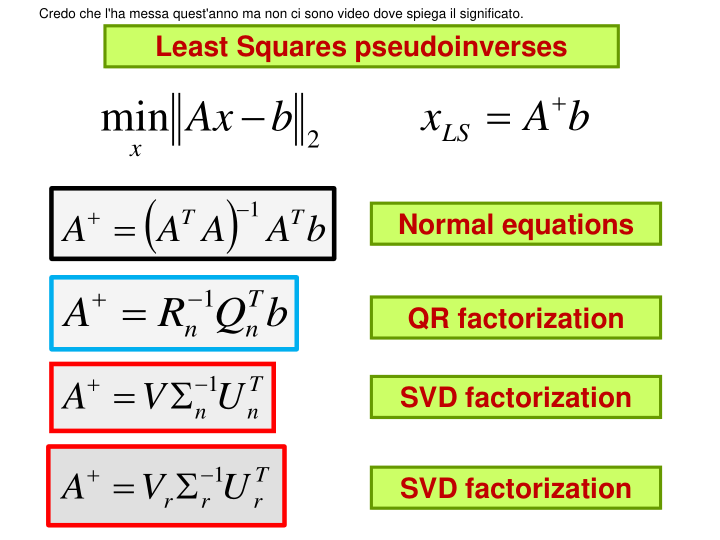
Quindi l'algoritmo è questo:  
calcolare la fatt ridsotta di svd  
calc il vettore c  
risolvere il sist diagonale  
ricostruisci x come V\*w



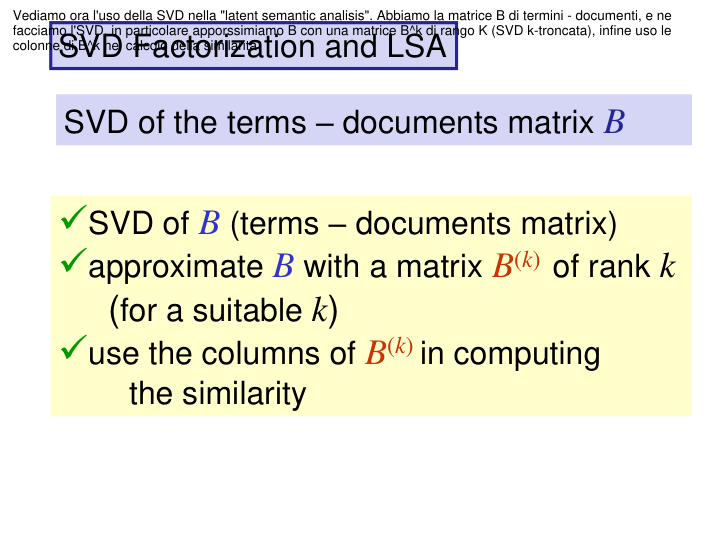
Ora afforntiamo il prob del minimi quadrati con il caso no full rank, qui si parla di Rank deficient perchè è minore di N. In questo problema esistono infinite Xls ma solo una ha lunghezza minima. Quindi procediamo in questo modo, A è = alla fattorizzazione economia con r, allora seguo la strada della proiezione ortogonale (Axls è la proiezione ortogonale) cioè faccio quella formuula UrUrtb, quindi invece di A metto la fattorizzazione (quel blocco in blu che è un sistema sottodeterminato), chiaramente si tratta di un sistema che ha infinite soluzioni, quello che si fa qui è moltiplicare ambo i membri prima per sigma^-1 e poi \* Vr così tolgo gli elementi vicino a xls. Otteniamo così la formula finale cioè il blocco rosso a sx. I lblocco a sx rosso invece è la risoluzioni che è quello che dobbiamo trovare., cioè quello di lunghezza minima. A+ è definita in quel modo quindi xls = A+ \* b



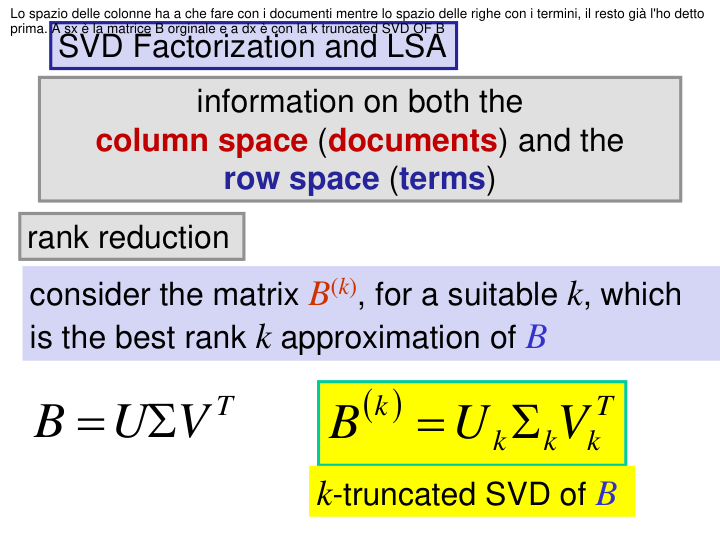
Credo che l'ha messa quest'anno ma non ci sono video dove spiega il significato.



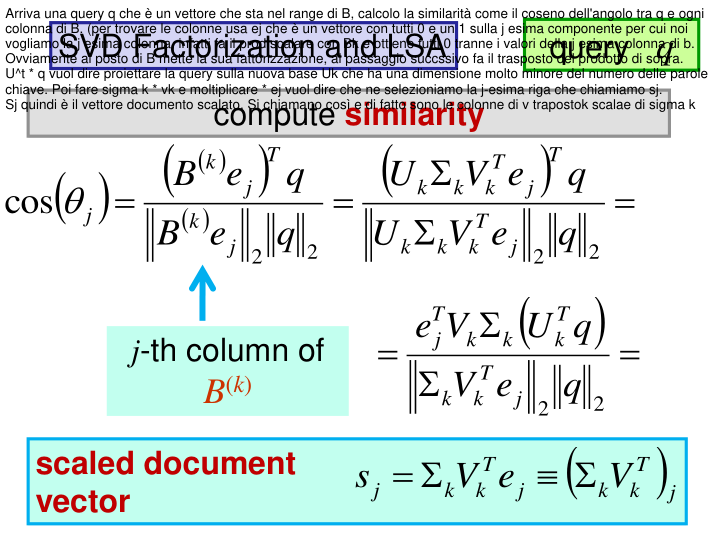
Vediamo ora l'uso della SVD nella "latent semantic analisis". Abbiamo la matrice B di termini - documenti, e ne facciamo l'SVD, in particolare apporssimiamo B con una matrice B^k di rango K (SVD k-troncata), infine uso le colonne di B^k nel calcolo della similarità



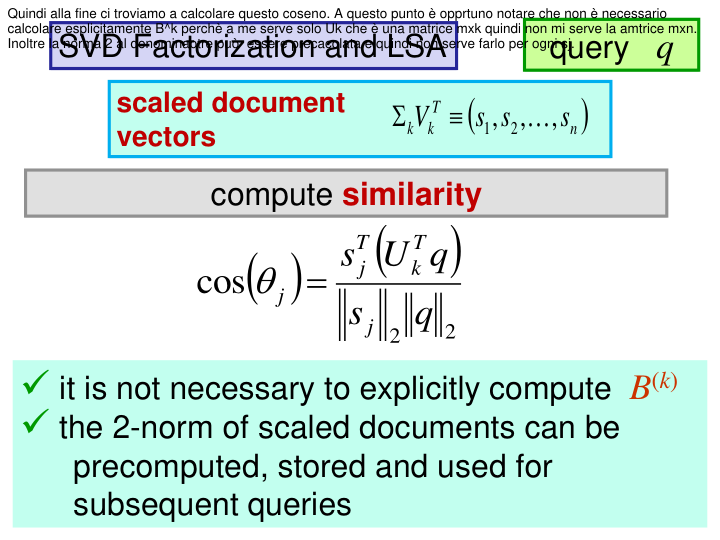
Lo spazio delle colonne ha a che fare con i documenti mentre lo spazio delle righe con i termini, il resto già l'ho detto prima. A sx è la matrice B orginale e a dx è con la k truncated SVD OF B



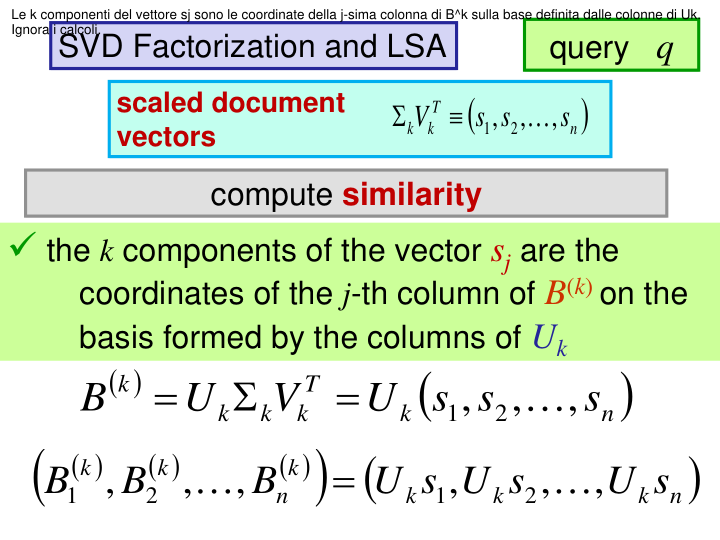
Arriva una query q che è un vettore che sta nel range di B, calcolo la similarità come il coseno dell'angolo tra q e ogni colonna di B. (per trovare le colonne usa ej che è un vettore con tutti 0 e un 1 sulla j esima componente per cui noi vogliamo la j esima colonna, infatti fa il prod scalare con Bk e ottiene tutti 0 tranne i valori della j esima colonna di b.  
Ovviamente al posto di B mette la sua fattorizzazione, al passaggio succssivo fa il trasposto del prodotto di sopra.  
U^t \* q vuol dire proiettare la query sulla nuova base Uk che ha una dimensione molto minore del numero delle parole chiave. Poi fare sigma k \* vk e moltiplicare \* ej vuol dire che ne selezioniamo la j-esima riga che chiamiamo sj.  
Sj quindi è il vettore documento scalato. Si chiamano così e di fatto sono le colonne di v trapostok scalae di sigma k



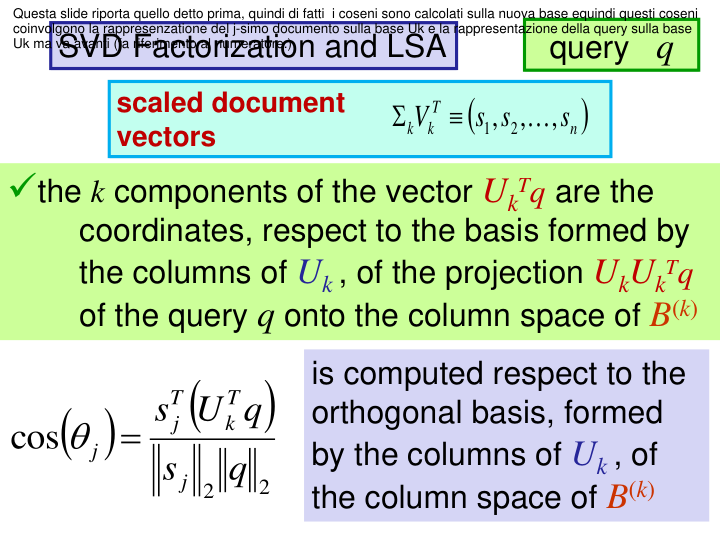
Quindi alla fine ci troviamo a calcolare questo coseno. A questo punto è opprtuno notare che non è necessario calcolare esplicitamente B^k perchè a me serve solo Uk che è una matrice mxk quindi non mi serve la amtrice mxn.  
Inoltre la norma 2 al denominaotre può essere precacolata e quindi non serve farlo per ogni sj.



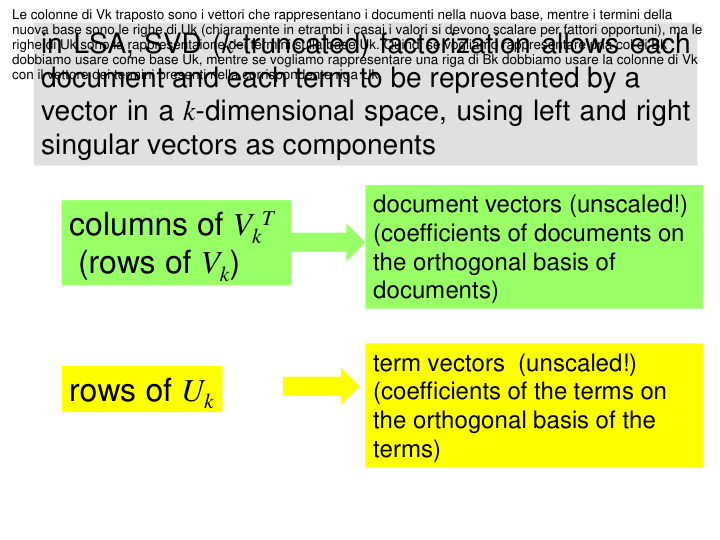
Le k componenti del vettore sj sono le coordinate della j-sima colonna di B^k sulla base definita dalle colonne di Uk.  
Ignora i calcoli.



Questa slide riporta quello detto prima, quindi di fatti i coseni sono calcolati sulla nuova base equindi questi coseni coinvolgono la rappresenzatione del j-simo documento sulla base Uk e la rappresentazione della query sulla base Uk ma va avanti (fa riferimento al numeratore.)



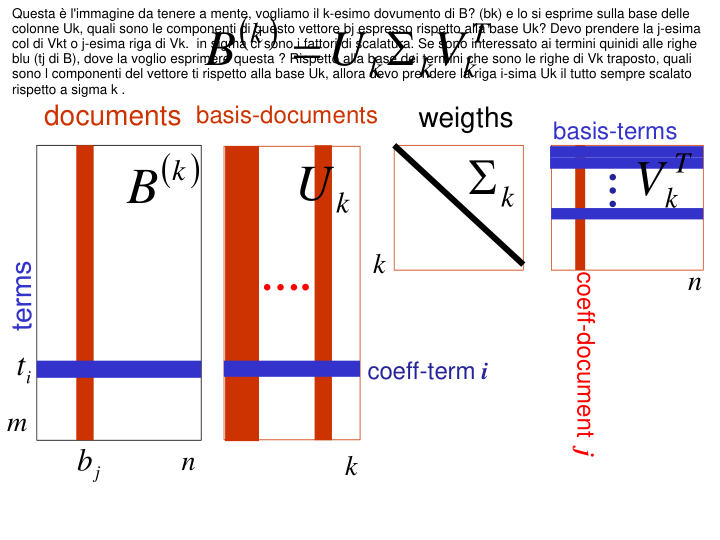
Le colonne di Vk traposto sono i vettori che rappresentano i documenti nella nuova base, mentre i termini della nuova base sono le righe di Uk (chiaramente in etrambi i casai i valori si devono scalare per fattori opportuni), ma le righe di Uk sono la rappresentaione dei termini sulla base Uk. Quindi se vogliamo rappresentare una col di Bk dobbiamo usare come base Uk, mentre se vogliamo rappresentare una riga di Bk dobbiamo usare la colonne di Vk con il vettore dei termini presenti nella corrispondente riga Uk



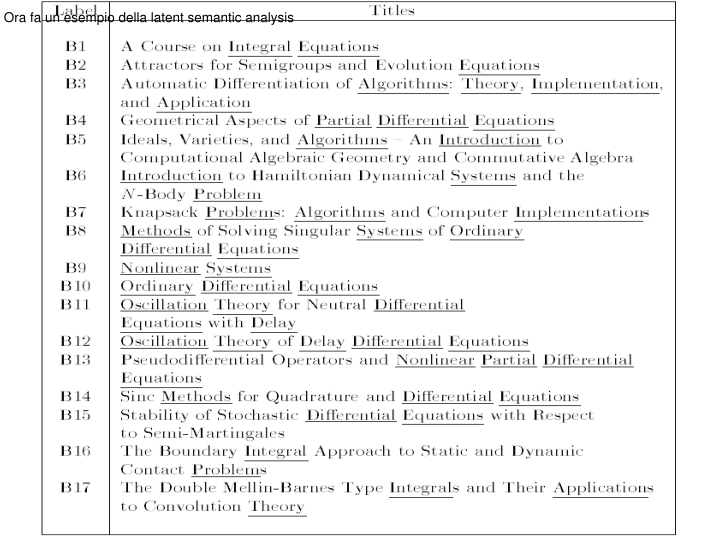
Qui c'è la trattazione che è la stessa slide che ha mostrato quando ha parlato della SVD completa, qui è riportata con la questione K. Alla fine otteniamo il vettore termini scalati, non aggiunge altro.



Questa è l'immagine da tenere a mente, vogliamo il k-esimo dovumento di B? (bk) e lo si esprime sulla base delle colonne Uk, quali sono le componenti di questo vettore bj espresso rispetto alla base Uk? Devo prendere la j-esima col di Vkt o j-esima riga di Vk. in sigma ci sono i fattori di scalatura. Se sono interessato ai termini quinidi alle righe blu (tj di B), dove la voglio esprimere questa ? Rispetto alla base dei termini che sono le righe di Vk traposto, quali sono l componenti del vettore ti rispetto alla base Uk, allora devo prendere la riga i-sima Uk il tutto sempre scalato rispetto a sigma k .

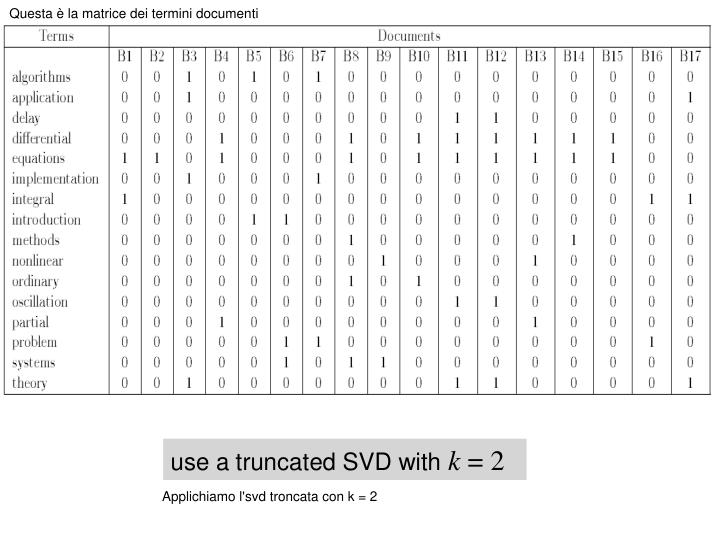


Ora fa un esempio della latent semantic analysis

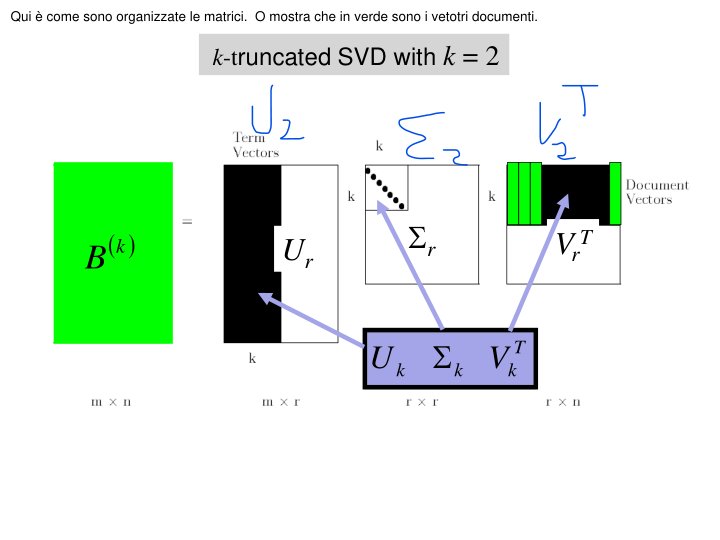


Questa è la matrice dei termini documenti

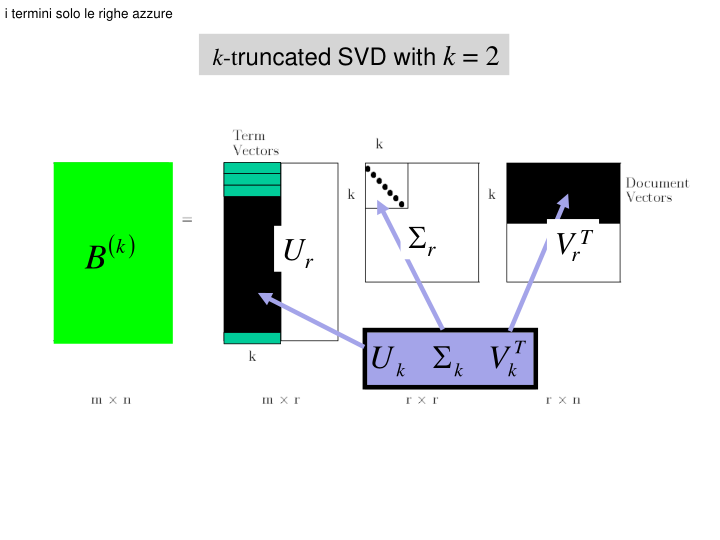
Applichiamo l'svd troncata con k = 2



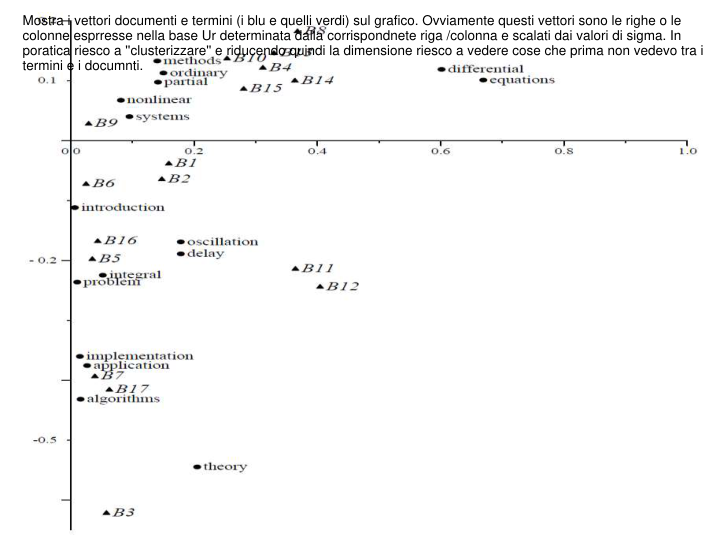
Qui è come sono organizzate le matrici. O mostra che in verde sono i vetotri documenti.



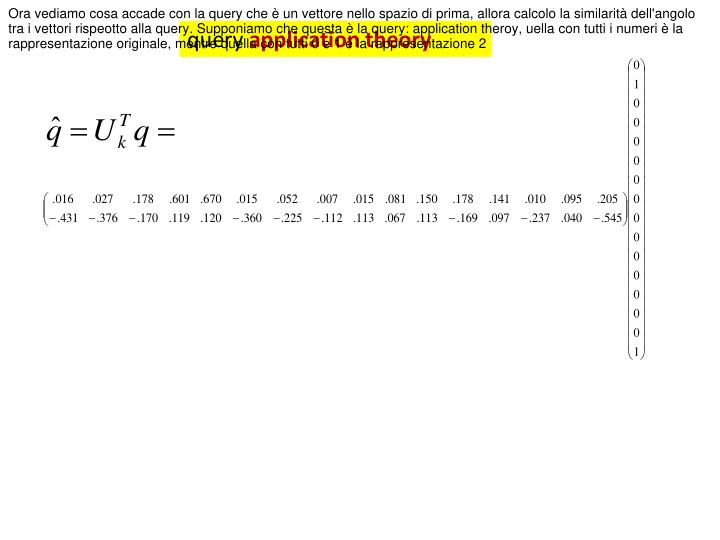
i termini solo le righe azzure



Mostra i vettori documenti e termini (i blu e quelli verdi) sul grafico. Ovviamente questi vettori sono le righe o le colonne esprresse nella base Ur determinata dalla corrispondnete riga /colonna e scalati dai valori di sigma. In poratica riesco a ''clusterizzare'' e riducendo quindi la dimensione riesco a vedere cose che prima non vedevo tra i termini e i documnti.

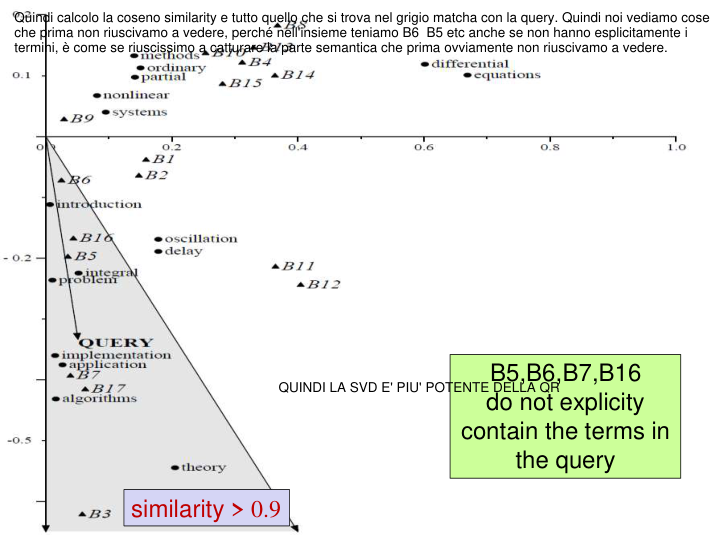


Ora vediamo cosa accade con la query che è un vettore nello spazio di prima, allora calcolo la similarità dell'angolo tra i vettori rispeotto alla query. Supponiamo che questa è la query: application theroy, uella con tutti i numeri è la rappresentazione originale, mentre quella con tutti 0 e 1 è la rappresentazione 2

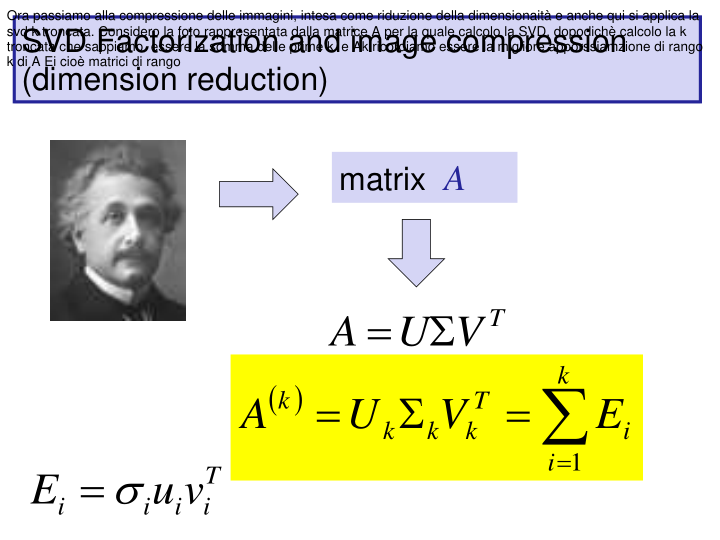


Quindi calcolo la coseno similarity e tutto quello che si trova nel grigio matcha con la query. Quindi noi vediamo cose che prima non riuscivamo a vedere, perché nell'insieme teniamo B6 B5 etc anche se non hanno esplicitamente i termini, è come se riuscissimo a catturare la parte semantica che prima ovviamente non riuscivamo a vedere.

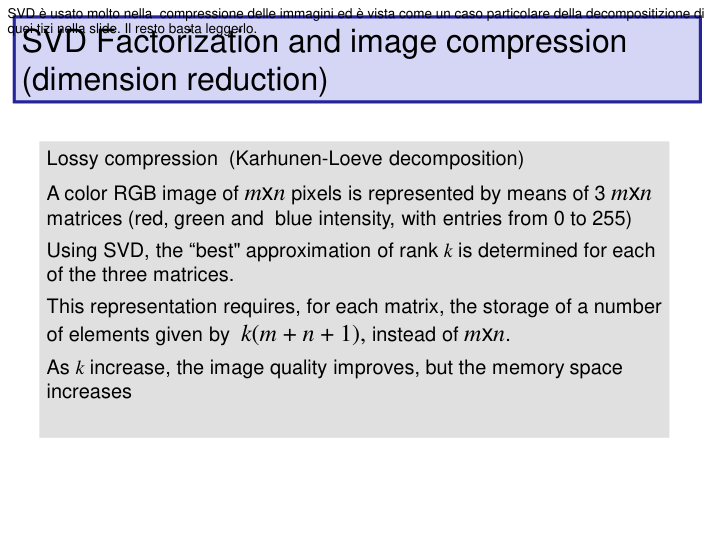
QUINDI LA SVD E' PIU' POTENTE DELLA QR



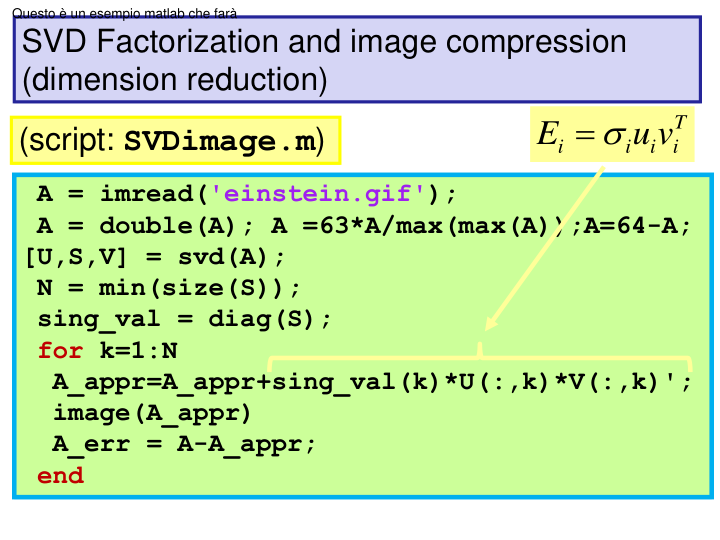
Ora passiamo alla compressione delle immagini, intesa come riduzione della dimensionaità e anche qui si applica la svd k troncata. Considero la foto rappresentata dalla matrice A per la quale calcolo la SVD, dopodichè calcolo la k troncata che sappiamo essere la somma delle prime k e Ak ricordiamo essere la migliore apporssiamzione di rango k di A Ei cioè matrici di rango



SVD è usato molto nella compressione delle immagini ed è vista come un caso particolare della decompositizione di quei tizi nella slide. Il resto basta leggerlo.

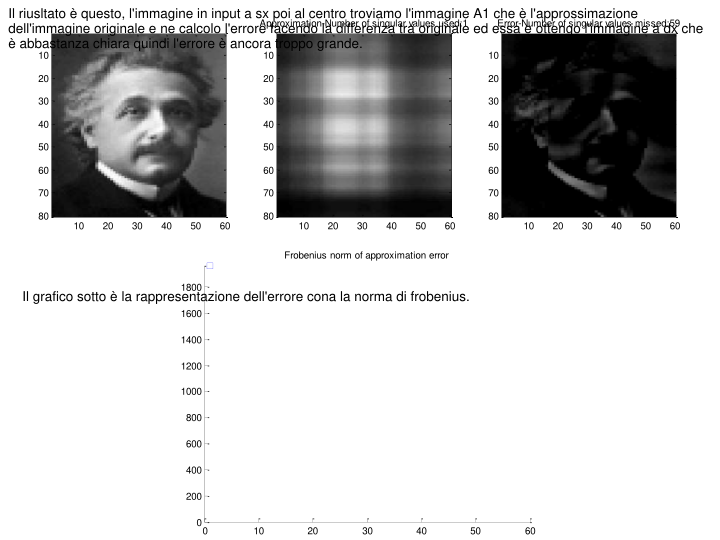


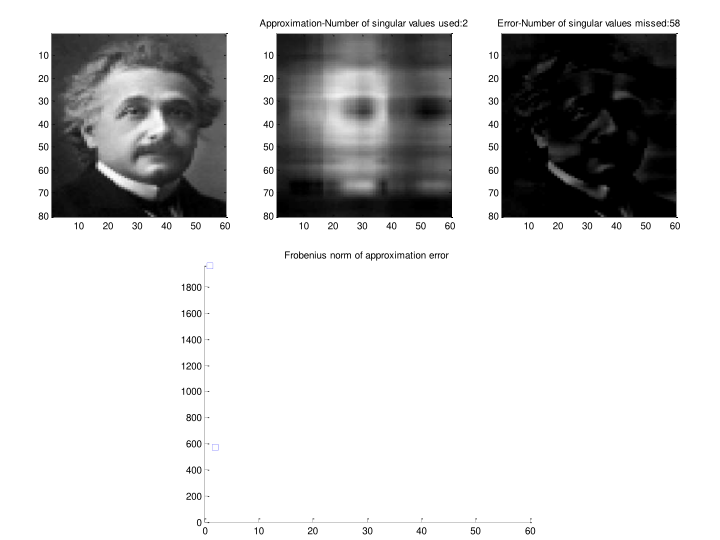
Questo è un esempio matlab che farà

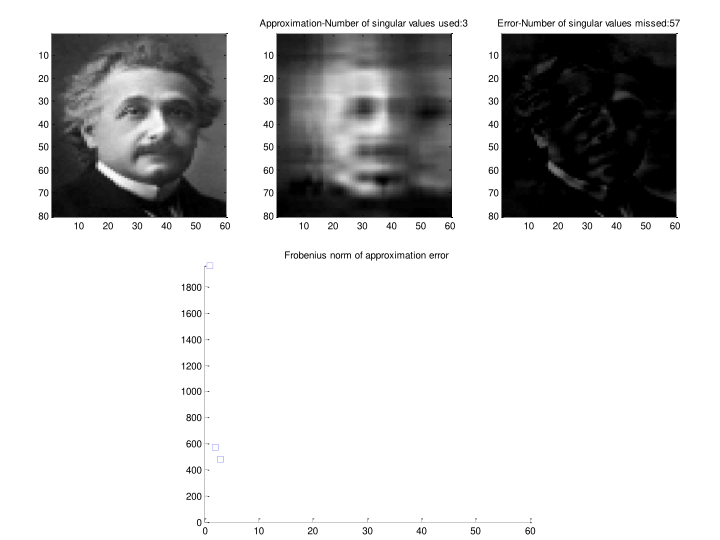


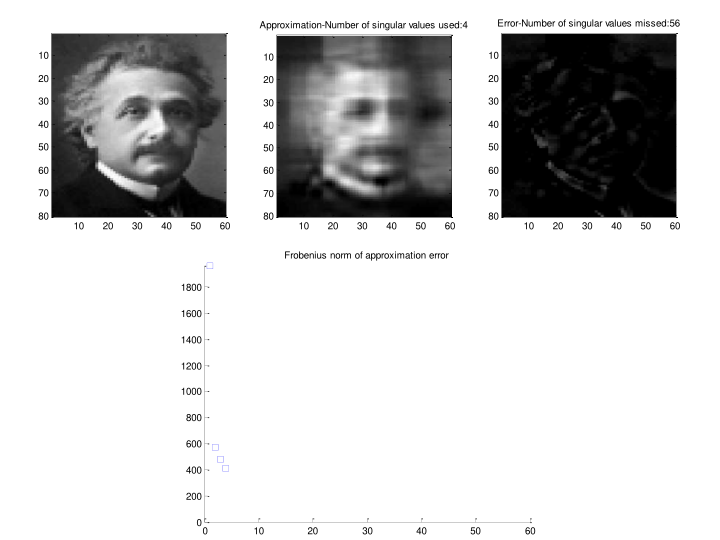
Il riusltato è questo, l'immagine in input a sx poi al centro troviamo l'immagine A1 che è l'approssimazione dell'immagine originale e ne calcolo l'errore facendo la differenza tra originale ed essa e ottengo l'immagine a dx che è abbastanza chiara quindi l'errore è ancora troppo grande.

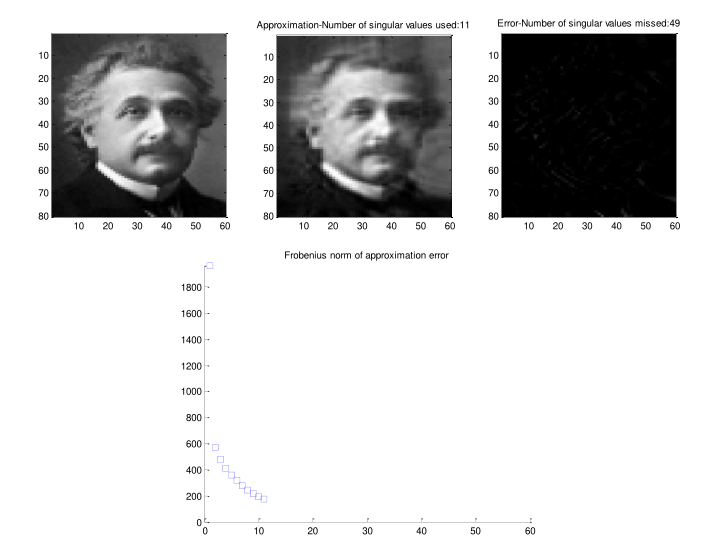
Il grafico sotto è la rappresentazione dell'errore cona la norma di frobenius.



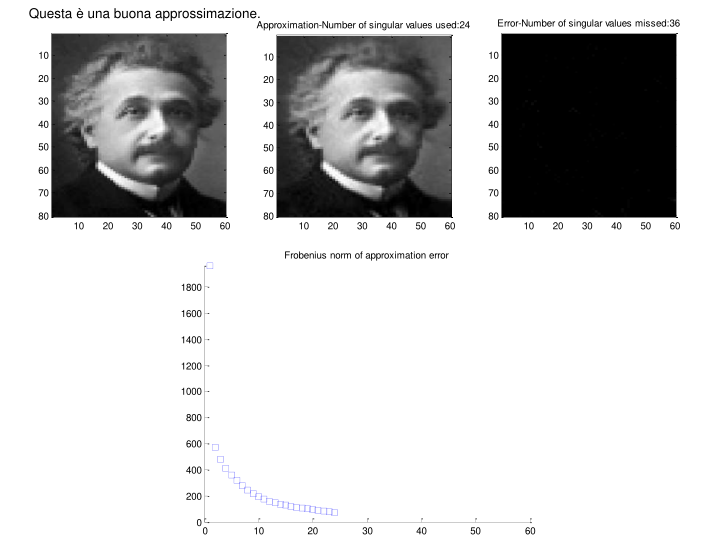




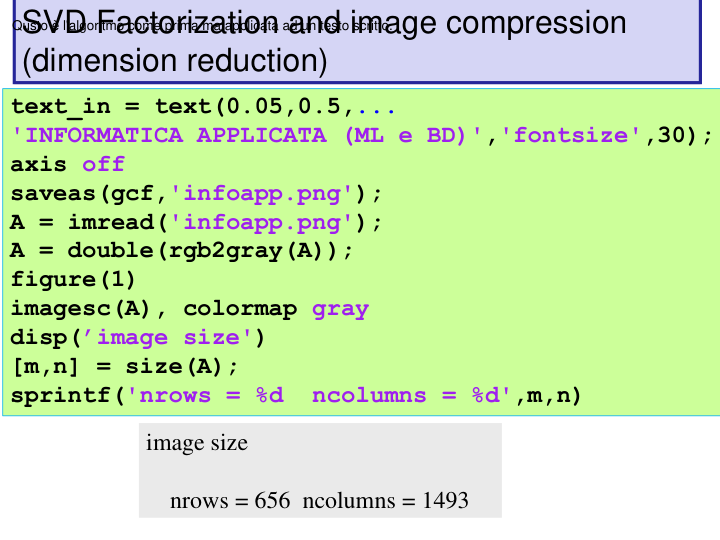


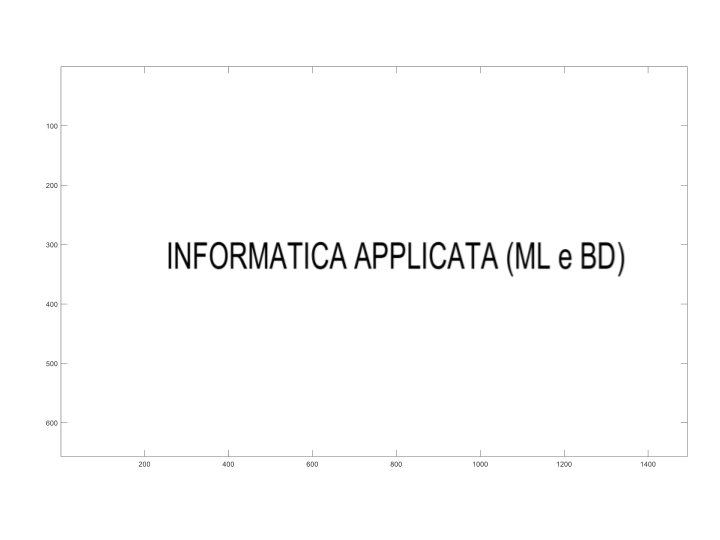


Questa è una buona approssimazione.

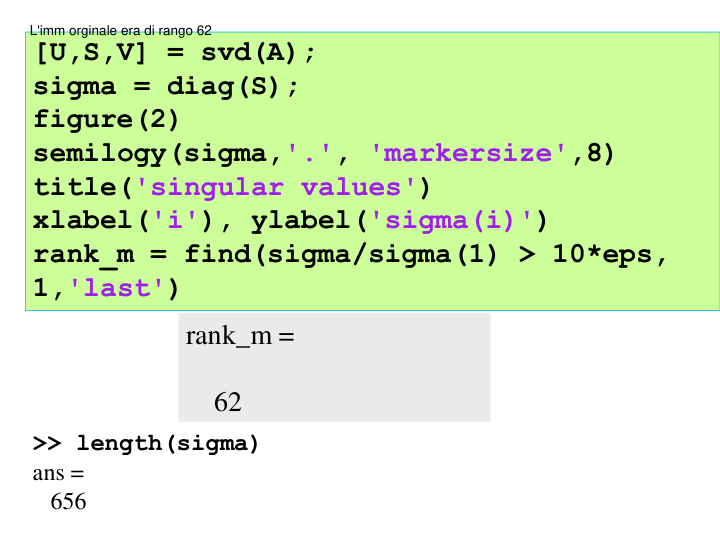


Qusto è l'algoritmo come prima ma applicata ad un testo scritto.

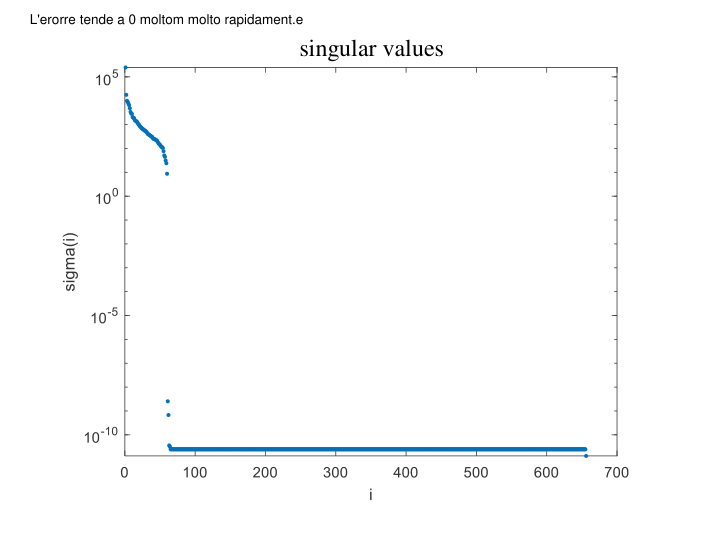


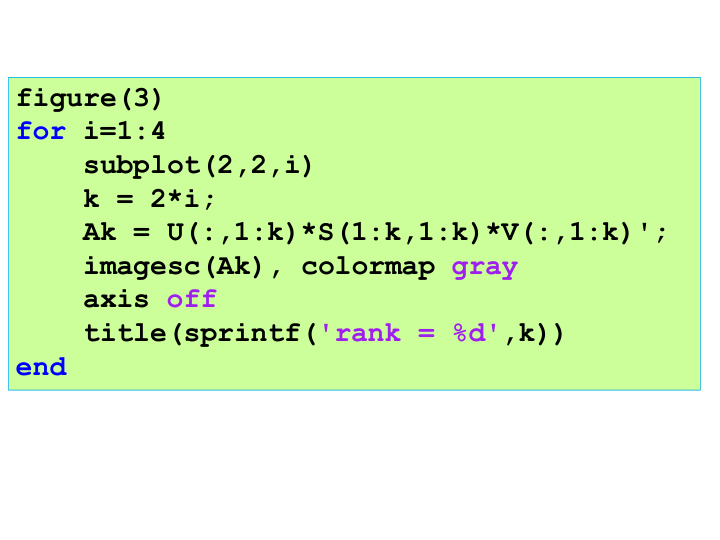


L'imm orginale era di rango 62



L'erorre tende a 0 moltom molto rapidament.e

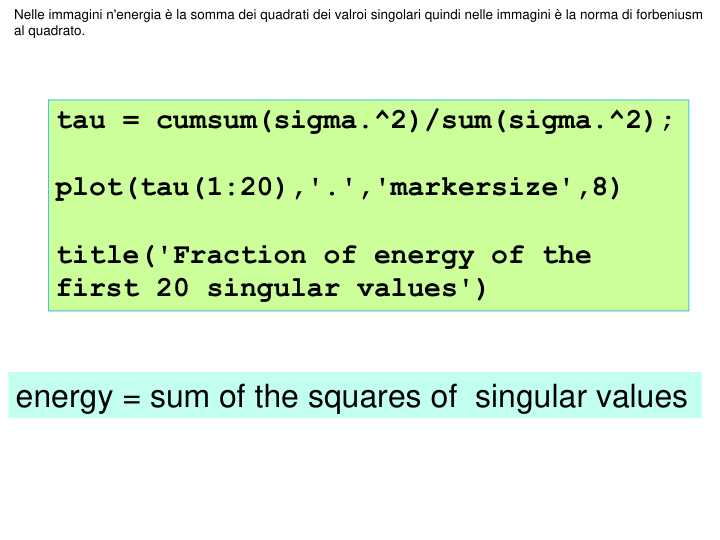


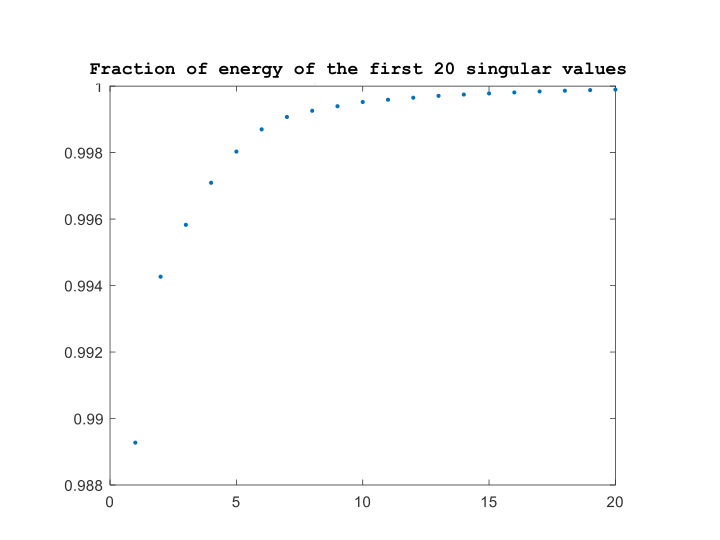


Già invece usando rango 4 riesco a capire ciò che è scritto.

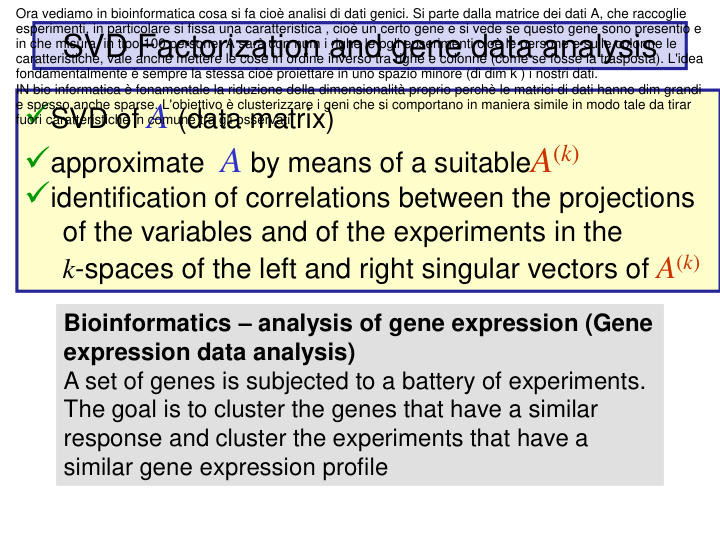


Nelle immagini n'energia è la somma dei quadrati dei valroi singolari quindi nelle immagini è la norma di forbeniusm al quadrato.





Ora vediamo in bioinformatica cosa si fa cioè analisi di dati genici. Si parte dalla matrice dei dati A, che raccoglie esperimenti, in particolare si fissa una caratteristica , cioè un certo gene e si vede se questo gene sono presentio e in che misura in tipo 100 persone. A sarà con num i righe le pgli epserimenti cioè le persone e sulle colonne le caratteristiche, vale anche mettere le cose in ordine inverso tra righe e colonne (come se fosse la trasposta). L'idea fondamentalmente è sempre la stessa cioè proiettare in uno spazio minore (di dim k ) i nostri dati.   
IN bio informatica è fonamentale la riduzione della dimensionalità proprio perchè le matrici di dati hanno dim grandi e spesso anche sparse. L'obiettivo è clusterizzare i geni che si comportano in maniera simile in modo tale da tirar fuori caratteristiche in comune tra gli osservati.



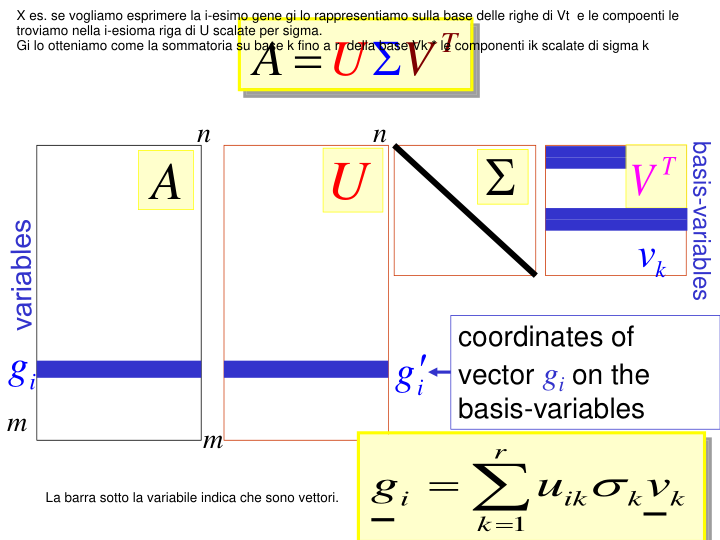
Come al solito questa è l'espressione SVD di A che ha la rappresentazione come in foto: voglio esprimere Aj sulla base delle colonne di U e le componenti sono a'j in Vt. Sigma ha il fattore di scala, questo ovviamente vale poi anche per le righe gi

Qui ha invertito mettnedo gli esperimento sulle colonne e le var sulle righe ma l'idea è la stessa.

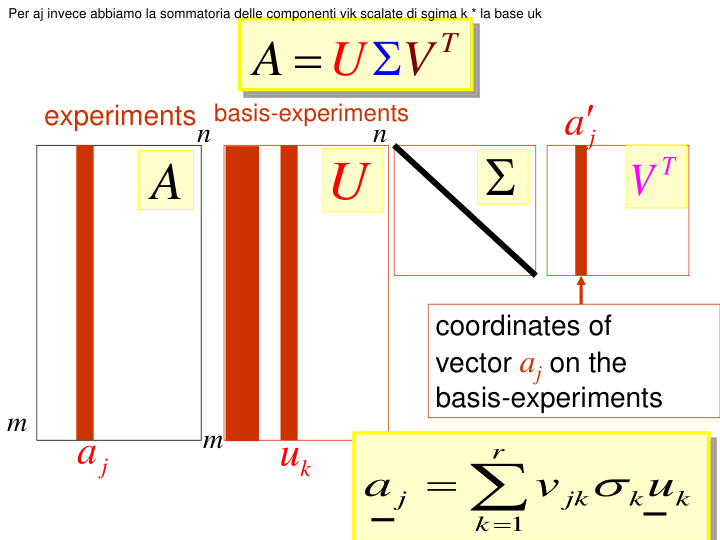


X es. se vogliamo esprimere la i-esimo gene gi lo rappresentiamo sulla base delle righe di Vt e le compoenti le troviamo nella i-esioma riga di U scalate per sigma.   
Gi lo otteniamo come la sommatoria su base k fino a r della base Vk \* le componenti ik scalate di sigma k

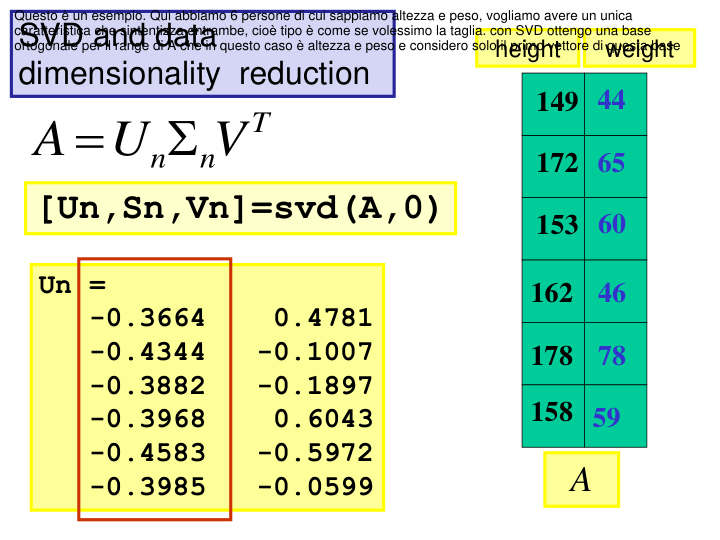
La barra sotto la variabile indica che sono vettori.



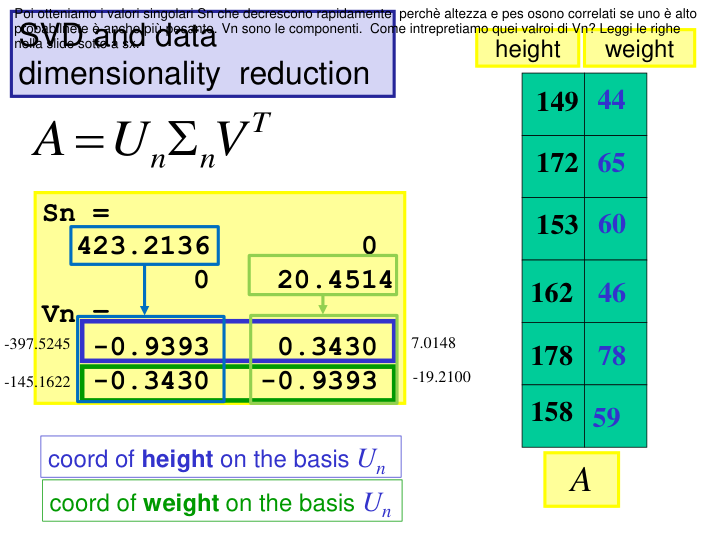
Per aj invece abbiamo la sommatoria delle componenti vik scalate di sgima k \* la base uk



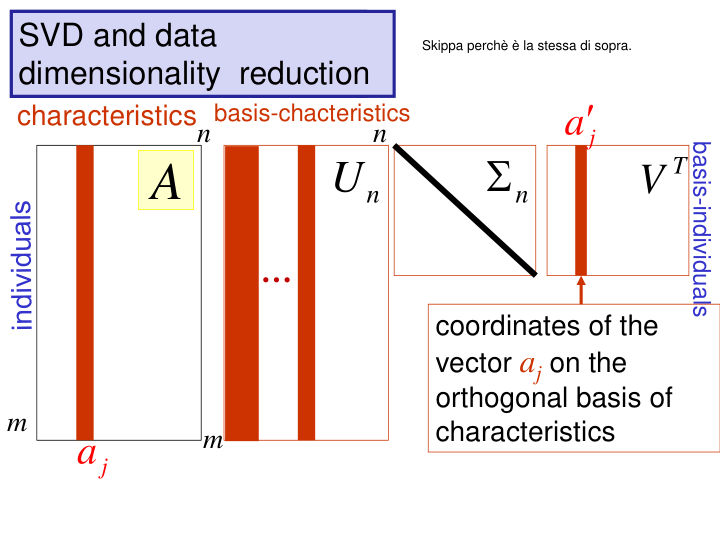
Questo è un esempio. Qui abbiamo 6 persone di cui sappiamo altezza e peso, vogliamo avere un unica caratteristica che sintentizza entrambe, cioè tipo è come se volessimo la taglia. con SVD ottengo una base ortogonale per il range di A che in questo caso è altezza e peso e considero solo il primo vettore di questa base



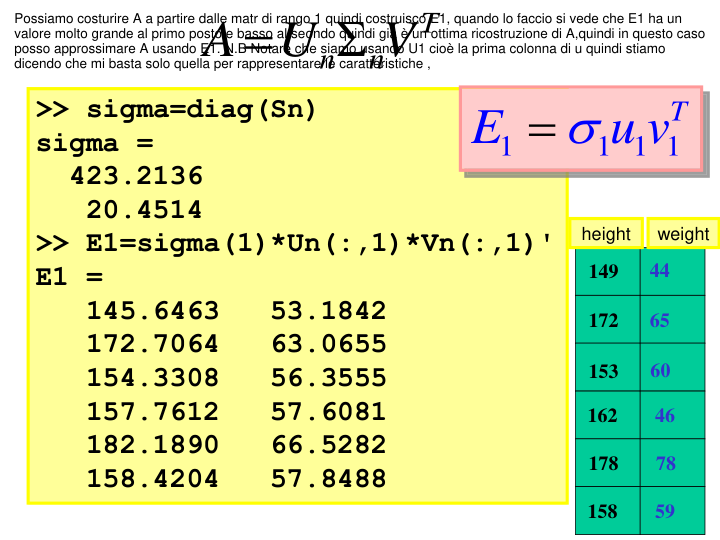
Poi otteniamo i valori singolari Sn che decrescono rapidamente perchè altezza e pes osono correlati se uno è alto probabilnete è anche più pesante. Vn sono le componenti. Come intrepretiamo quei valroi di Vn? Leggi le righe nella slide sotto a sx.



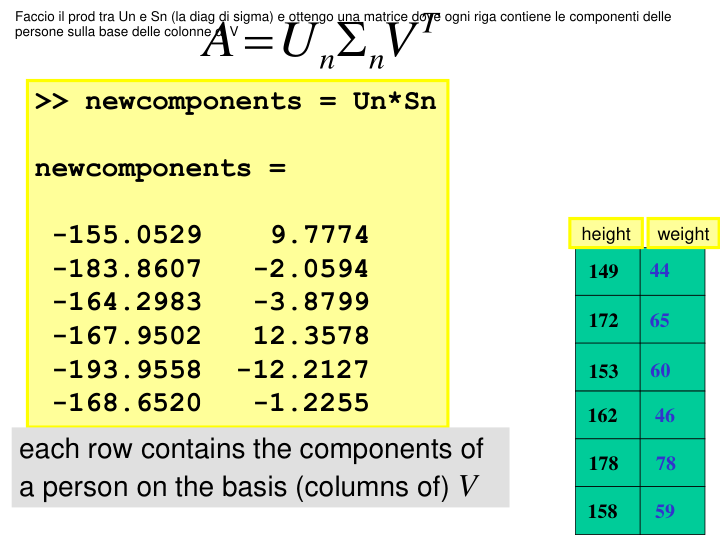
Skippa perchè è la stessa di sopra.



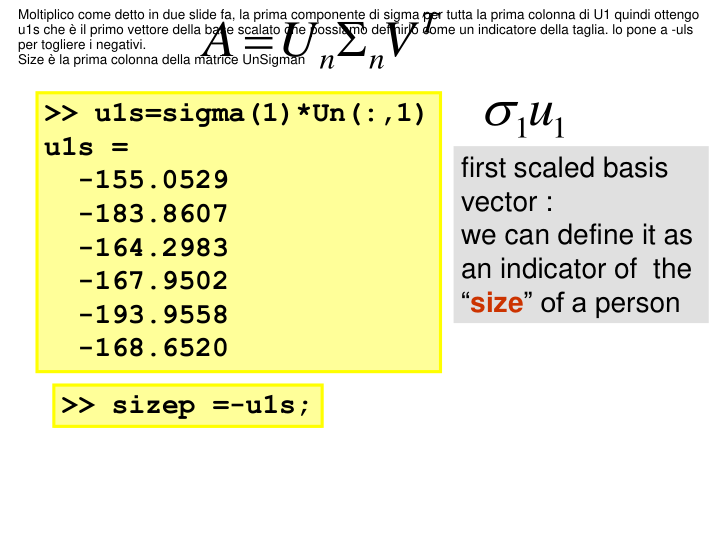
Possiamo costurire A a partire dalle matr di rango 1 quindi costruisco E1, quando lo faccio si vede che E1 ha un valore molto grande al primo posto e basso al seondo quindi già è un ottima ricostruzione di A,quindi in questo caso posso approssimare A usando E1. N.B Notare che siamo usando U1 cioè la prima colonna di u quindi stiamo dicendo che mi basta solo quella per rappresentare le caratteristiche ,



Faccio il prod tra Un e Sn (la diag di sigma) e ottengo una matrice dove ogni riga contiene le componenti delle persone sulla base delle colonne di V

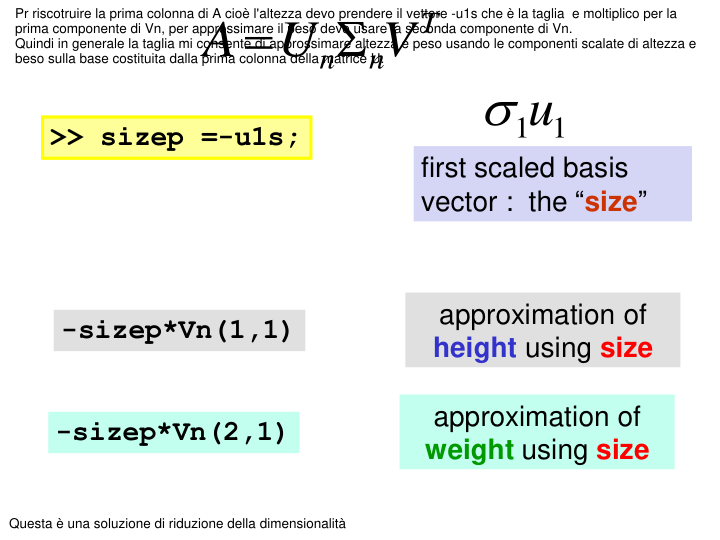


Moltiplico come detto in due slide fa, la prima componente di sigma per tutta la prima colonna di U1 quindi ottengo u1s che è il primo vettore della base scalato che possiamo definirlo come un indicatore della taglia. lo pone a -uls per togliere i negativi.  
Size è la prima colonna della matrice UnSigman

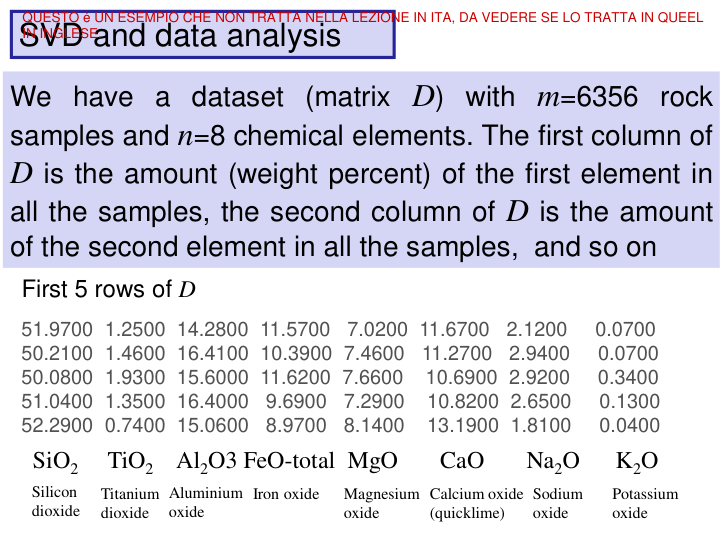


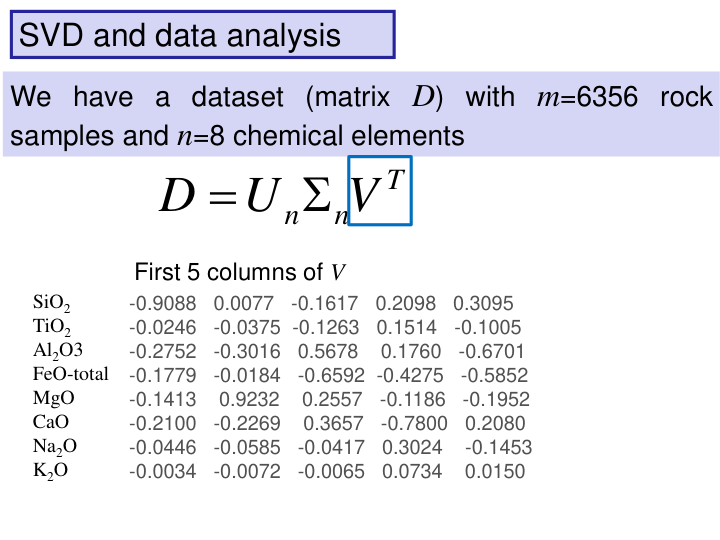
Pr riscotruire la prima colonna di A cioè l'altezza devo prendere il vettore -u1s che è la taglia e moltiplico per la prima componente di Vn, per approssimare il peso devo usare la seconda componente di Vn.  
Quindi in generale la taglia mi consente di approssimare altezza e peso usando le componenti scalate di altezza e beso sulla base costituita dalla prima colonna della matrice U.

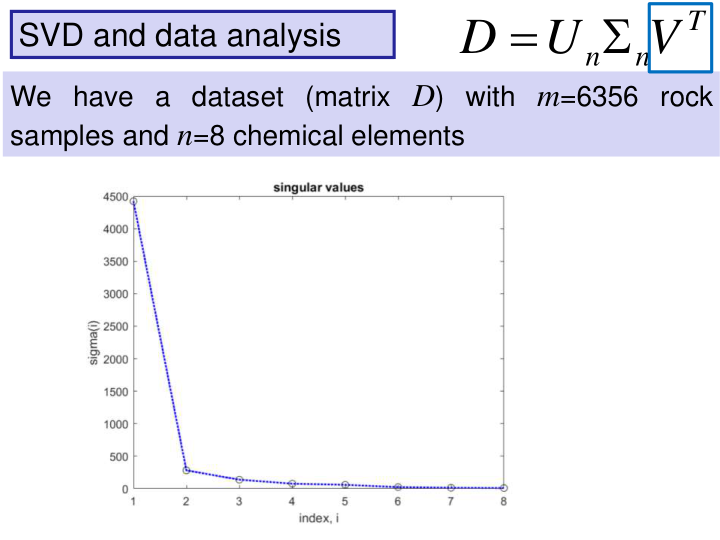
Questa è una soluzione di riduzione della dimensionalità

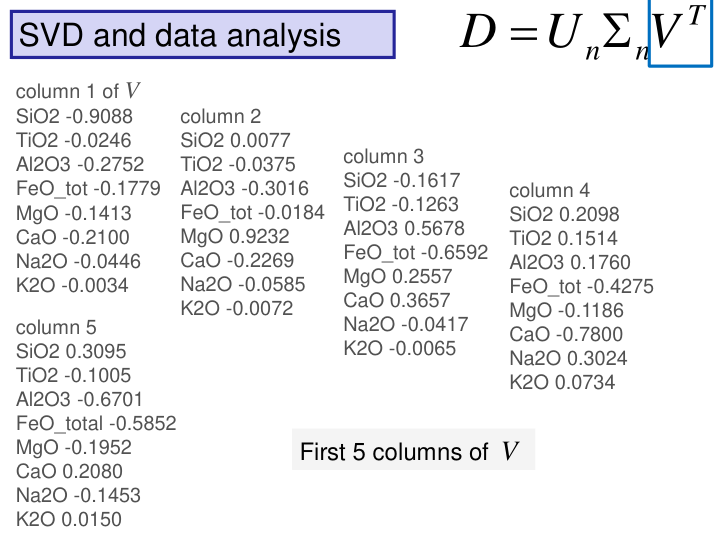


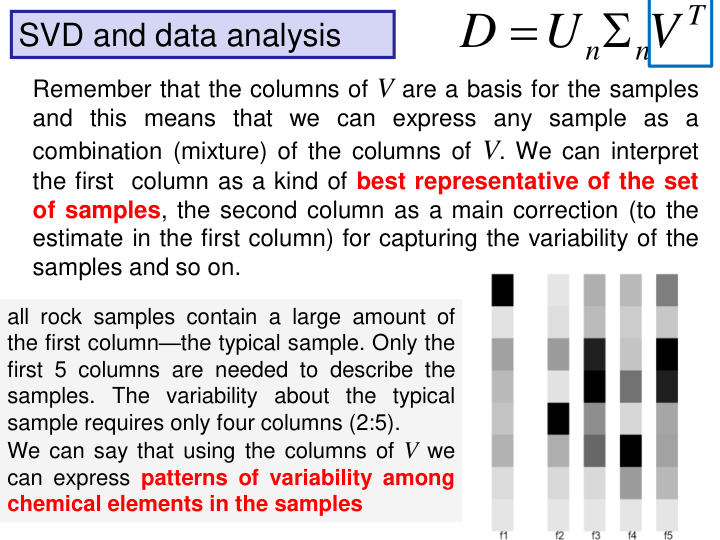
QUESTO è UN ESEMPIO CHE NON TRATTA NELLA LEZIONE IN ITA, DA VEDERE SE LO TRATTA IN QUEEL IN INGLESE.

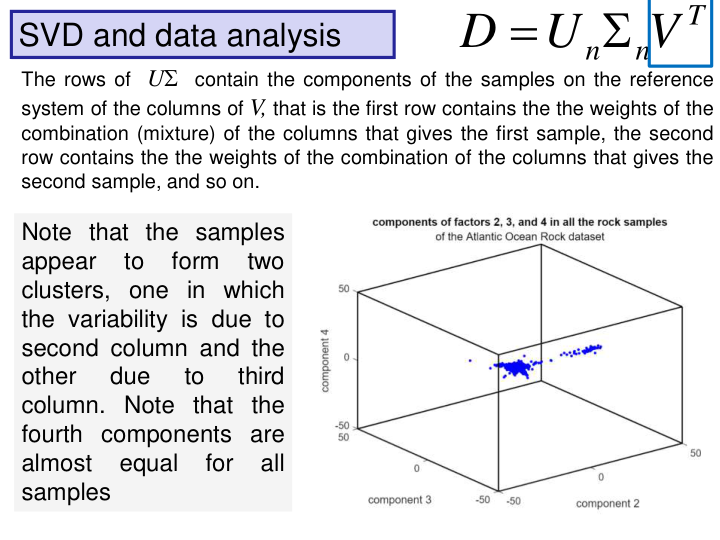


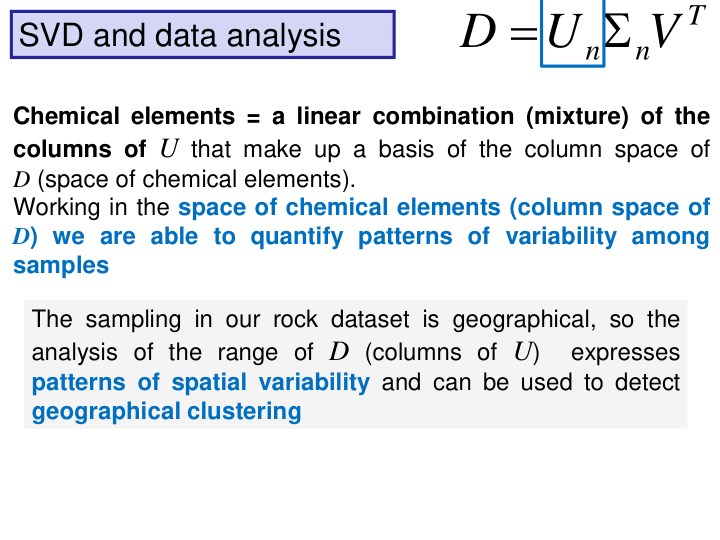


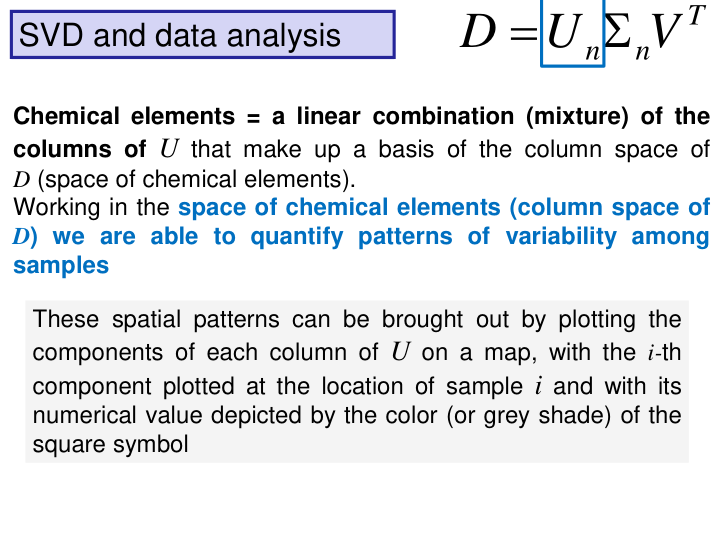


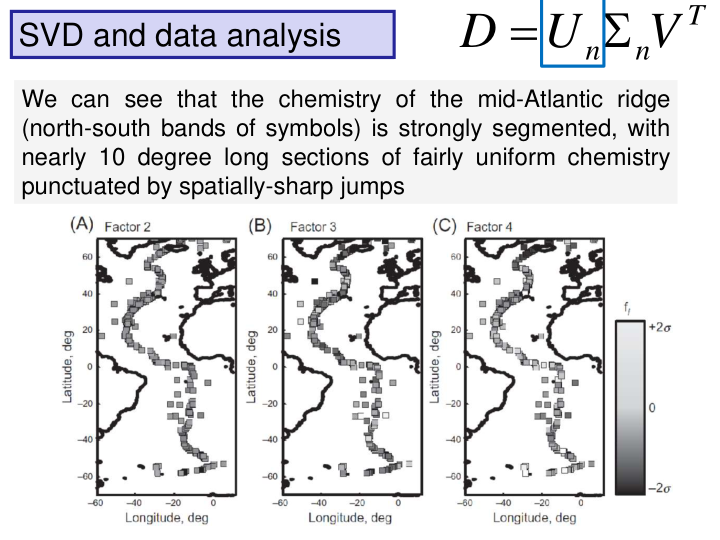












https://web.microsoftstream.com/video/2b8ecd39-0c13-4b7a-9082-3c10e9782517

