Feb 8 2007

⟨ Singular Value Decomposition ⟩

- Why do we need this method?
  - Solving linear algebraic eq.
  - least squares fit
  - Image processing or noisy signal filtering

- What can this method do?

Solve $Ax=b$ when A is nonsingular
$M=N$ — A is singular

- (a) when many solutions exist, give a particular solution
  $Ax=0$ → construct a general solution
- (b) when no solution exists, do LSF

$M<N$ general solutions

$M>N$ LSF

- SVD theorem

(To construct U, V matrices, we need to know how to diagonalize or find eigenvalues and eigenvectors of matrices)

→ so here I will explain the method and you will use subroutines provided

- Examples

Feb 7, 2006

⟨Singular Value Decomposition⟩

- This method will diagnose for you precisely what the problem is when Gaussian elimination and LU decomposition fail to give satisfactory results.

- This method will be also used for solving most linear least squares problems.

- SVD Theorem: Any $M \times N$ matrix A (→ singular or nonsingular) can be expressed as

$M \geq N$

$$\begin{array}{c} \\ M \end{array}\overbrace{\begin{pmatrix} \\ A \\ \\ \end{pmatrix}}^{N} = \begin{pmatrix} \\ U \\ \\ \end{pmatrix} \overbrace{\begin{pmatrix} w_1 & & \\ & w_2 & \\ & & \ddots \\ & & & w_N \end{pmatrix}}^{W} \begin{pmatrix} \\ V^T \\ \\ \end{pmatrix} \quad \cdots \text{①}$$

transpose

$V^T = V^{-1}$ $\quad U^T = U^{-1}$

U : $M \times N$ column-orthogonal matrix

W : $N \times N$ diagonal matrix with positive or zero elements

V : $N \times N$ orthogonal matrix.

Columns are orthonormal

$$\sum_{i=1}^{M} U_{ik} U_{in} = \delta_{kn} \qquad 1 \leq k, n \leq N$$

$$\sum_{j=1}^{N} V_{jk} V_{jn} = \delta_{kn} \qquad 1 \leq k, n \leq N$$

$$\sum_{j=1}^{N} V_{kj} V_{nj} = \delta_{kn}$$

$V V^T = I$

$$\begin{pmatrix} U^T \end{pmatrix} \begin{pmatrix} U \end{pmatrix} = I, \quad \begin{pmatrix} V^T \end{pmatrix} \begin{pmatrix} V \end{pmatrix} = I$$

$$Ax=b \qquad x=A^{-1}b = (UWV^T)^{-1}b$$
$$\Rightarrow \boxed{x=(VW^{-1}U^T)b}$$

- The decomposition ① can always be done no matter how singular the matrix is.

- The decomposition is unique up to (i) making the same permutation of the columns of U, elements of W, and columns of V (or rows of V^T), or (ii) forming linear combinations of any columns of U and V whose corresponding elements of W happen to be exactly equal.

- Algorithm of singular value decomposition
  (How to find U, V, W matrices)

  (i) Find the eigenvalues of the matrix A^TA and arrange them in descending order. $\{\lambda_i,\ i=1,\cdots,N\}$

  (ii) Find the number of nonzero eigenvalues of the matrix A^TA. Set it to r.

  (iii) Find the orthonormal eigenvectors $\{v_i,\ i=1,\cdots,N\}$ of the matrix A^TA corresponding to the obtained eigenvalues, and arrange them in the same order to form the column-vectors of the matrix V.

  (iv) Form a diagonal matrix W placing on the leading diagonal of it the square roots of the eigenvalues of the matrix A^TA in descending order

  $$w_i=\sqrt{\lambda_i}$$

(V) Find the first column-vectors of the matrix U from $u_i = w_i^{-1} A v_i \quad (i=1,\cdots,r)$

~~(VI) Add to the matrix U the rest of $m-r$ vectors using the Gram-Schmidt orthogonalization method.~~

unnecessary or ~~(V)* Find the eigenvalues and the corresponding orthonormal eigenvectors $\{u_i,\ i=1,\cdots,N'\}$ of the matrix AA^T, and arrange the eigenvectors such that the eigenvalues descend, to form the column-vectors of the matrix U.~~

※ Remarks : (I) Why is the matrix V (U) constructed from the orthonormal eigenvectors of the matrix A^TA (AA^T)

$$A^TA = (U W V^T)^T (U W V^T)$$

$$= (V W^T U^T) U W V^T$$

$$\xrightarrow{U^TU = I} \; = V(W^T W)V^T$$

$$\boxed{\underbrace{W^TW}_{\text{diagonal matrix}} = V^T(A^TA)V}$$

From the similarity transformation,

$$\underbrace{\tilde{C}}_{\text{diagonal matrix}} = R^T \underbrace{C}_{\text{general matrix}} R$$

$R \rightarrow$ consists of column-vectors made of orthonormal eigenvectors of C

$$V = \Big(\underbrace{\boxed{}}_{\text{eigenvector}} \; \boxed{} \; \boxed{} \; \cdots \; \boxed{} \Big)$$

In analogy with the similarity transformation,

V is formed from the orthonormal eigenvectors of A^TA.

W : Its diagonal elements consisty of the square roots of the eigenvalues of the matrix A^TA.

$$AA^T = (UWV^T)(UWV^T)^T$$
$$= UWV^TVW^TU^T$$
$$= U(WW^T)U^T$$

$\Rightarrow$ U : formed from the column vectors made of the orthonormal eigenvectors of AA^T.

e.g. : Find the singular value decomposition of the matrix A

: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\underset{\text{row column}}{3 \times 2}$ matrix $\quad M=3, \; N=2$

① $A^TA = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0 \Rightarrow (2-\lambda)^2 - 1 = 0$$
$$\lambda^2 - 4\lambda + 3 = 0$$
$$\lambda_1 = 3, \; \lambda_2 = 1$$

② # of nonzero eigenvalues of the matrix A^TA
$= r = 2$

③ Find the orthonormal eigenvectors of A^TA.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = 3\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

$2x_1 + y_1 = 3x_1$ $\qquad x_1 = y_1 \quad \Rightarrow \quad V_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

$x_1 + 2y_1 = 3y_1$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

$x_1 = -y_1 \quad \Rightarrow \quad V_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$

$$\boxed{V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}}$$

④ Find the diagonal matrix W: $w_i = \sqrt{\lambda_i}$

$$\underset{2\times 2}{W} = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \qquad \underset{3\times 2}{W} = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

⑤* $$AA^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

or
$$u_1 = w_1^{-1} A v_1$$
$$u_2 = w_2^{-1} A v_2$$
$$u_3 = e_1 - (u_1^T e_1) u_1 - (u_2^T e_1) u_2$$

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (2-\lambda)(1-\lambda)^2 - 2(1-\lambda) = 0$$

$$\lambda_1 = 3, \quad \lambda_2 = 1, \quad \lambda_3 = 0$$

$$u_1 = \begin{pmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{pmatrix} \qquad u_2 = \begin{pmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \qquad u_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix}$$

constructed from eigenvectors w/ nonzero eigenvalues.

$$\Rightarrow \text{So } A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{3} & 0 \\ \frac{\sqrt{6}}{6} & -\frac{1}{\sqrt{2}} \\ \frac{\sqrt{6}}{6} & \frac{1}{\sqrt{2}} \end{pmatrix} \times \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

or

$$A = \begin{pmatrix} \frac{\sqrt{6}}{3} & 0 & \frac{1}{\sqrt{3}} \\ \frac{\sqrt{6}}{6} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{\sqrt{6}}{6} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

e.g. Another example: Find the singular value decomposition of the matrix $A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix}$

Case 1: $M = N$ (A : square matrix)

$\Rightarrow U, W, V$: square matrices

$$Ax = b$$

$$x = A^{-1} b$$

$$= (U W V^T)^{-1} b$$

$$= V W^{-1} U^T b$$

$$= V \cdot \left[\mathrm{diag}\left(\frac{1}{w_j} \right) \right] U^T b$$

w_j → diagonal elements of W

A^{-1} does not exist if w_j is very small or zero

(A is singular)

$\Rightarrow$ SVD gives a clear diagnosis of the problem!!

- condition number of a matrix:

$$K(A) = \frac{|\lambda|_{max}}{|\lambda|_{min}} \qquad \lambda : \text{eigenvalue of } A$$

in our case

$$K(A) = \frac{|w_j|_{max}}{|w_j|_{min}}$$

A is singular if $K(A)$ is infinite

A is "ill-conditioned" if $K(A)$ is too large

(If $K^{-1}(A) \approx 10^{-12}$ for double precision)

Let me review the concepts of "rank" and "nullity".

Let A be an m by n matrix.

row-rank of A : maximum # of linearly independent vectors of the row-vectors

column-rank of A : maximum # of linearly independent vectors of the column-vectors

rank = column-rank

nullity = dimension of null space of A
: # of free variables in the solution of $AX=0$.

- Rank theorem:

$$\text{rank}(A) + \text{nullity}(A) = n = \text{\# of columns}$$

- e.g. Compute the rank and nullity of the matrix

$$A = \begin{bmatrix} 2 & -4 & 0 & 2 & 1 \\ -1 & 2 & 1 & 2 & 3 \\ 1 & -2 & 1 & 4 & 4 \end{bmatrix}$$

$AX = b$

do the same linear combination to b vector.

Using the Gaussian elimination,

the second row + first row $\times \left(\frac{1}{2}\right)$

$= 0 \quad 0 \quad 1 \quad 3 \quad \frac{7}{2}$

the third row + first row $\times \left(-\frac{1}{2}\right)$

$= 0 \quad 0 \quad 1 \quad 3 \quad \frac{7}{2}$

$$\Rightarrow \begin{bmatrix} 2 & -4 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & \frac{7}{2} \\ 0 & 0 & 1 & 3 & \frac{7}{2} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 2 & -4 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 & \frac{7}{2} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

column-rank = maximum # of linearly independent column vectors = 2

basis : $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

nullity : 3

because there are three free variables that satisfy $Ax=0$.

$$2x_1 - 4x_2 + 2x_4 + x_5 = 0$$
$$x_3 + 3x_4 + \frac{7}{2}x_5 = 0$$

rank(A) + nullity(A) = 5 !!

* For any matrix A (square or rectangular or real or complex),

column rank = row rank = rank(A)

$\uparrow$

# of linearly independent column vectors

rank(A) = rank(A^T)

* Definition of the <u>nullspace</u> of a matrix A

: set of x that satisfies $Ax = 0$.

(dimension of the nullspace = nullity)

* Definition of the <u>range</u> of A

: set of b for which $Ax = b$ has a solution for x.

(dimension of the range = rank)

rank + nullity = N

dimension of the matrix

Now, for a square matrix A ($M=N$)

SVD of $A = UWV^T$ can be written as

$A v_j = w_j u_j$ where u_j and v_j are columns of U and V

$j = 1, \cdots, N$.

- $w_j = 0$: v_j is in the nullspace of A.
- $w_j \neq 0$: u_j is in the range of A

because when $w_j \neq 0$ v_j is in the orthogonal complement of the nullspace.

$\Leftrightarrow$ v_j is in the solution space : $Ax=b$

e.g. $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ $\det(A) = 0$

A^{-1} does not exist

so A is singular!!

$\boxed{\begin{array}{l} \text{rank} = 1 \\ \text{nullity} = 1 \end{array}}$

$A^TA = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix}$, $AA^T = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}$

eigenvalues of A^TA and AA^T : $\lambda_1 = 10$, $\lambda_2 = 0$

$W = \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix}$

$$U = \begin{array}{c} \begin{array}{cc} w_1=\sqrt{10} & w_2=0 \end{array} \\ \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix} \end{array} \qquad V = \begin{array}{c} \begin{array}{cc} w_1=\sqrt{10} & w_2=0 \end{array} \\ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \end{array}$$

$x_1 + x_2 = 0$

$Ax = 0$

basis for the range of A

basis for the orthogonal complement of the range

basis for the nullspace

basis for the orthogonal complement of the nullspace(A)

$$A = UWV^T = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \boxed{\sqrt{10}} & 0 \\ 0 & \boxed{0} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T$$

orthonormal basis for the range

orthormal basis for the nullspace

$M=N$, A: singular ($Ax=b$)

$\boxed{1}$ If b lies in the range of A:

SVD can give a particular solution.

singular set of equations does have a solution x

since rank $< N$, there are fewer equations (linearly independent) than unknowns.

$\Rightarrow \exists$ more than one solutions because any vector in the nullspace ($Ax=0$) can be added to x in any combination. This would be a perfectly good solution as well.

To pick out a solution with the smallest length $|x|^2$, we may use SVD.

Replace $\frac{1}{\omega_j}$ by zero if $\omega_j = 0$.

Then compute $x = V \cdot \left[\mathrm{diag}\left(\frac{1}{\omega_j}\right) \right] \left(U^T \cdot b \right)$.

(proof) Consider $|\vec{x}+\vec{x}'|$, where $\vec{x}'$ lies in the nullspace.
Let $\widetilde{W}^{-1}$ be the modified inverse of W with some elements zeroed.

$$|\vec{x}+\vec{x}'| = \left|\left(V\cdot\widetilde{W}^{-1}U^T b\right) + \vec{x}'\right|$$

$$= \left|V\left(\widetilde{W}^{-1}U^T b + V^T\vec{x}'\right)\right|$$

$$= \left|\underbrace{\widetilde{W}^{-1}U^T b} + \underbrace{V^T\vec{x}'}\right|$$

since the columns of V are orthonormal.

first term: nonzero j components only where $w_j \neq 0$ since b is in the range.
second term: nonzero j components only when $w_j = 0$ since $\vec{x}'$ is in the nullspace.

So the minimum length is obtained for $\vec{x}'=0$.

e.g. From the previous example,

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \qquad Ax=b \Leftrightarrow \begin{matrix} x_1+y_1=b_1 \\ 2(x_1+y_1)=b_2 \end{matrix}$$

from U, basis set for the range

$$= \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$$

$\Rightarrow$ If $b_2 = 2b_1$, $\vec{b}$ is in the range of A.

To obtain a solution with the smallest length,

modified inverse of W with some elements zeroed.

$$W = \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix}$$

Using $x = V \tilde{W}^{-1} U^T b$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{b_1}{5\sqrt{2}} + \frac{2b_2}{5\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{b_1}{10} + \frac{2b_2}{10} \\ \frac{b_1}{10} + \frac{2b_2}{10} \end{bmatrix}$$

$$x_1 = \frac{1}{10}(b_1 + 2b_2)$$

$$y_1 = \frac{1}{10}(b_1 + 2b_2)$$

condition for existence of many solutions

$Ax = 0$

plug $b_2 = 2b_1$ into the above.

$x_1 = y_1 = \frac{b_1}{2}$ $\Leftarrow$ This is what SVD provides.

General solution: $\begin{matrix} x_1 = \\ y_1 = \end{matrix} \begin{pmatrix} \frac{b_1}{2} \\ \frac{b_1}{2} \end{pmatrix} + \alpha \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$, $\alpha = \text{const}$

Compare this with the direct solution from $Ax = b$.

$$x_1 + y_1 = b_1$$

linear combination of columns of V corresponding to $w_j = 0$ eigenvalues.

$\boxed{2}$ If b does <u>not</u> lie in the range of A;
the set of equations $(Ax=b)$ has <u>no</u> solution.

BUT, SVD will provide the closest possible values that satisfy the equations in the least squares sense.

$$\Rightarrow \left(\begin{array}{l} x \text{ that minimizes } |Ax-b|. \\ \Leftrightarrow x = V \cdot \tilde{W}^{-1} U^T b \end{array} \right)$$

(proof) Suppose that we modify x by adding some arbitrary x'.

$$Ax-b \longrightarrow A(x+x')-b = Ax-b+b', \quad \text{where } b' = Ax'$$

b' is in the range.

$$|Ax-b+b'| = |\overbrace{(UWV^T)}^{A}\overbrace{(V\tilde{W}^{-1}U^T b)}^{x} - b + b'|$$

$$= |(UW\tilde{W}^{-1}U^T - 1)b + b'|$$

$$= |U \cdot [(W\tilde{W}^{-1} - 1) \cdot U^T \cdot b + U^T b']|$$

$$= |\underbrace{(W \cdot \tilde{W}^{-1} - 1) \cdot U^T b}_{\substack{\text{nonzero for } w_j = 0 \\ \text{because of } (W\tilde{W}^{-1} - 1)}} + \underbrace{U^T \cdot b'}_{\text{nonzero for } w_j \neq 0}|$$

So the minimum can be obtained when $b' = 0$.

e.g. From the previous example,

$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ $\quad Ax=b \iff \begin{array}{l} x_1+y_1 = b_1 \\ 2(x_1+y_1) = b_2 \end{array}$

If $b_2 \neq 2b_1$, there are no solutions for $Ax=b$.
[$\vec{b}$ is not in the range.]

SVD provides $\left.\begin{array}{l} x_1 = \frac{1}{10}(b_1+2b_2) \\ y_1 = \frac{1}{10}(b_1+2b_2) \end{array}\right\}$.

Least-squares fit:

$$f = (x_1+y_1-b_1)^2 + (2x_1+2y_1-b_2)^2$$

$$0 = \frac{\partial f}{\partial x_1} = 2(x_1+y_1-b_1) + 2(2x_1+2y_1-b_2)\cdot 2$$

$$0 = \frac{\partial f}{\partial y_1} = 2(x_1+y_1-b_1) + 2(2x_1+2y_1-b_2)\cdot 2$$

$$\Rightarrow 5x_1+5y_1-b_1-2b_2 = 0$$

$$\left. x_1+y_1 = \frac{b_1+2b_2}{5} \right\}$$

$\Rightarrow$ SVD agrees w/ LSF.

- Rule of thumb :
  - Make small w_j's zero first
  - Then Use $x = V \tilde{W}^{-1} U^T b$

Case 2 : $M < N$ (fewer equations than unknowns)

$Ax = b$

There is no unique solution.

There will be an $N-M$ dimensional family of solutions.

SVD can provide the whole solution space for you.

How?

(i) Augment the matrix A with rows of zeros underneath its M nonzero rows, until it is filled up to be square, $N \times N$.

(ii) Augment the b vetor similarly with zeros to make it $N \times 1$ vector.

Now you have an $N \times N$ singular matrix A.

(iii) Do SVD. → obtain a particular solution x_p

[Make sure to find small w_j's and to zero them before performing SVD.]

(iv) solution space : $x = x_p + \{ \sum_i \alpha_i V_i \}$

linear combination of columns of V corresponding to zeroed w_j's.

e.g.

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix} \qquad \begin{matrix} M=2 \\ N=3 \end{matrix}$$

$$Ax = b \iff \begin{matrix} 3x_1 + x_2 + x_3 = b_1 \\ -x_1 + 3x_2 + x_3 = b_2 \end{matrix}$$

(i) Augment the matrix A

$$\tilde{A} = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

(ii) Augment the b vector : $\tilde{b} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}$

(iii) Do SVD.

$$A^T A = \begin{pmatrix} 3 & -1 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{pmatrix}$$

$$AA^T = \begin{pmatrix} 11 & 1 & 0 \\ 1 & 11 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Eigenvalues of AA^T and A^TA :

$\lambda_1 = 12$, $\lambda_2 = 10$, $\lambda_3 = 0$

$$W = \begin{bmatrix} \sqrt{12} & 0 & 0 \\ 0 & \sqrt{10} & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \tilde{W}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{12}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 & \boxed{0} \end{bmatrix}$$

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & -\frac{5}{\sqrt{30}} \end{bmatrix}$$

basis vector corresponds to $w_j = 0$

particular solution

$$\rightarrow X_p = V \tilde{W}^{-1} U^T b$$

$$= \begin{bmatrix} \frac{17}{60} & \frac{-7}{60} & 0 \\ \frac{4}{60} & \frac{16}{60} & 0 \\ \frac{1}{12} & \frac{1}{12} & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

$$X_1 = \frac{17}{60} b_1 - \frac{7}{60} b_2$$

$$X_2 = \frac{b_1}{15} + \frac{4}{15} b_2$$

$$X_3 = \frac{1}{12}(b_1 + b_2)$$

general solution:

$$X = \begin{pmatrix} \frac{17}{60}b_1 - \frac{7}{60}b_2 \\ \frac{b_1}{15} + \frac{4}{15}b_2 \\ \frac{b_1}{12} + \frac{b_2}{12} \end{pmatrix} + \alpha \begin{pmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ -\frac{5}{\sqrt{30}} \end{pmatrix}$$

α = (undetermined) constant.

Check if the general solution satisfies the original equations $Ax = b$.

Case 3: $M > N$ (More equations than unknowns)
$Ax = b$

SVD will find x that minimizes $|Ax - b|$.

$$x = V\tilde{W}^{-1}U^T b$$

e.g. From the previous example,

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \tilde{W}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$A = 3\times2$

$V: 2\times2$ $\quad \tilde{W}^{-1}: 2\times3$

$U = 3\times3$

$x = 2\times1$

$b = 3\times1$

$$\tilde{W}^{-1}\tilde{W} = I$$

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Using the reduced SVD
(as in the Numerical recipe)

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tilde{W}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix}$$

$$V \tilde{W}^{-1} U^T = \overset{2\times 2}{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}} \overset{2\times 2}{\begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix}} \overset{2\times 3}{\begin{pmatrix} \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}}$$

$$= \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

$V\tilde{W}^{-1}U^T$

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

$x = V\tilde{W}^{-1}U^T b$ $\qquad x : 2\times 1 \,,\; b : 3\times 1$

$$\boxed{\begin{aligned} x_1 &= \tfrac{1}{3}b_1 - \tfrac{1}{3}b_2 + \tfrac{2}{3}b_3 \\ y_1 &= \tfrac{1}{3}b_1 + \tfrac{2}{3}b_2 - \tfrac{1}{3}b_3 \end{aligned}}$$

Least-square fit: $\quad Ax = b$

$$f \equiv (x_1 + y_1 - b_1)^2 + (y_1 - b_2)^2 + (x_1 - b_3)^2$$

$$0 = \frac{\partial f}{\partial x_1} = 2(x_1 + y_1 - b_1) + 2(x_1 - b_3) \Rightarrow \quad 2x_1 + y_1 = b_1 + b_3$$

$$0 = \frac{\partial f}{\partial y_1} = 2(x_1 + y_1 - b_1) + 2(y_1 - b_2) \Rightarrow \quad x_1 + 2y_1 = b_1 + b_2$$

$$\boxed{\begin{aligned} y_1 &= \tfrac{1}{3}(b_1 + 2b_2 - b_3) \\ x_1 &= \tfrac{1}{3}(b_1 - b_2 + 2b_3) \end{aligned}}$$

< Application of SVD >

① Solving linear algebraic equations

② Linear least squares fit

③ Image processing
or noisy signal filtering

$\overset{SVD}{\Rightarrow}$

$$A_{ij} = \sum_{l=1}^{N} w_l U_{il} V_{jl} \qquad i=1,\dots,M \quad j=1,\dots,N$$

→ your noise component

If most of w_j's are very small
(only a small number K of w_j's are significant), then
we can approximate the sum into

your signal component.

$$A^{(K)}_{ij} = \sum_{l=1}^{K} w_l U_{il} V_{jl} \qquad i=1,\dots,M \quad j=1,\dots,N$$

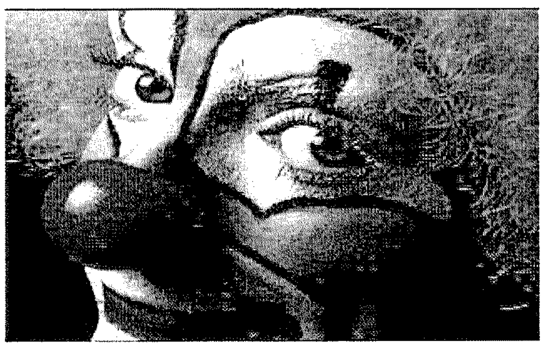
⇒ To compute $A \cdot x$, one needs $K(M+N)$ multiplications instead of MN for the full matrix.

Application: Image Compression

- View $m \times n$ image as a (real) matrix A, find best rank K approx. by SVD
- Storage $K(m+n)$ instead of mn

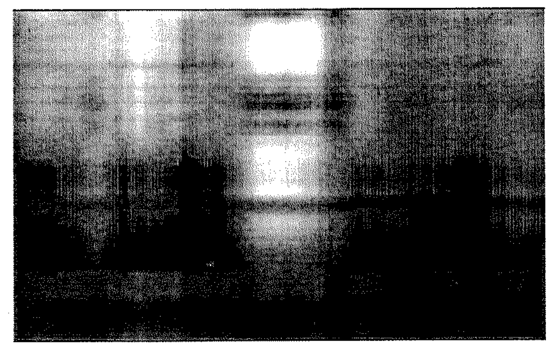
K

Original (Rank 200)

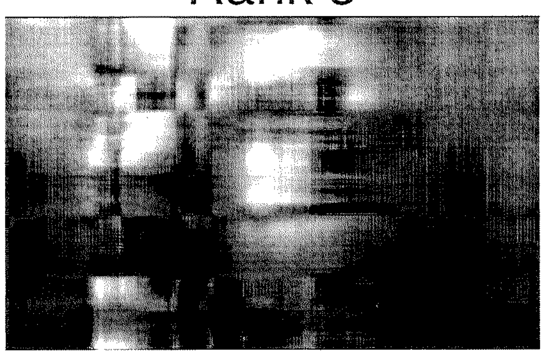


Rank 1

Rank 2



Rank 5



Rank 15

Rank 50