SVD:
Singular Value Decomposition

Motivation

$A:$ any matrix
with complete set of e-vectors
$A = S\Lambda S^{-1}$
$A^{-1} = S\Lambda^{-1}S^{-1}$
$x = S\Lambda^{-1}S^{-1}b$

$A:$ any matrix
$A = LU$
$LUx = b$
$Ly = b$
$Ux = y$

Assume A full rank

$A:$ *symmetric* matrix
$A = Q\Lambda Q^T$
$A^{-1} = Q\Lambda^{-1}Q^T$
$x = Q\Lambda^{-1}Q^T b$

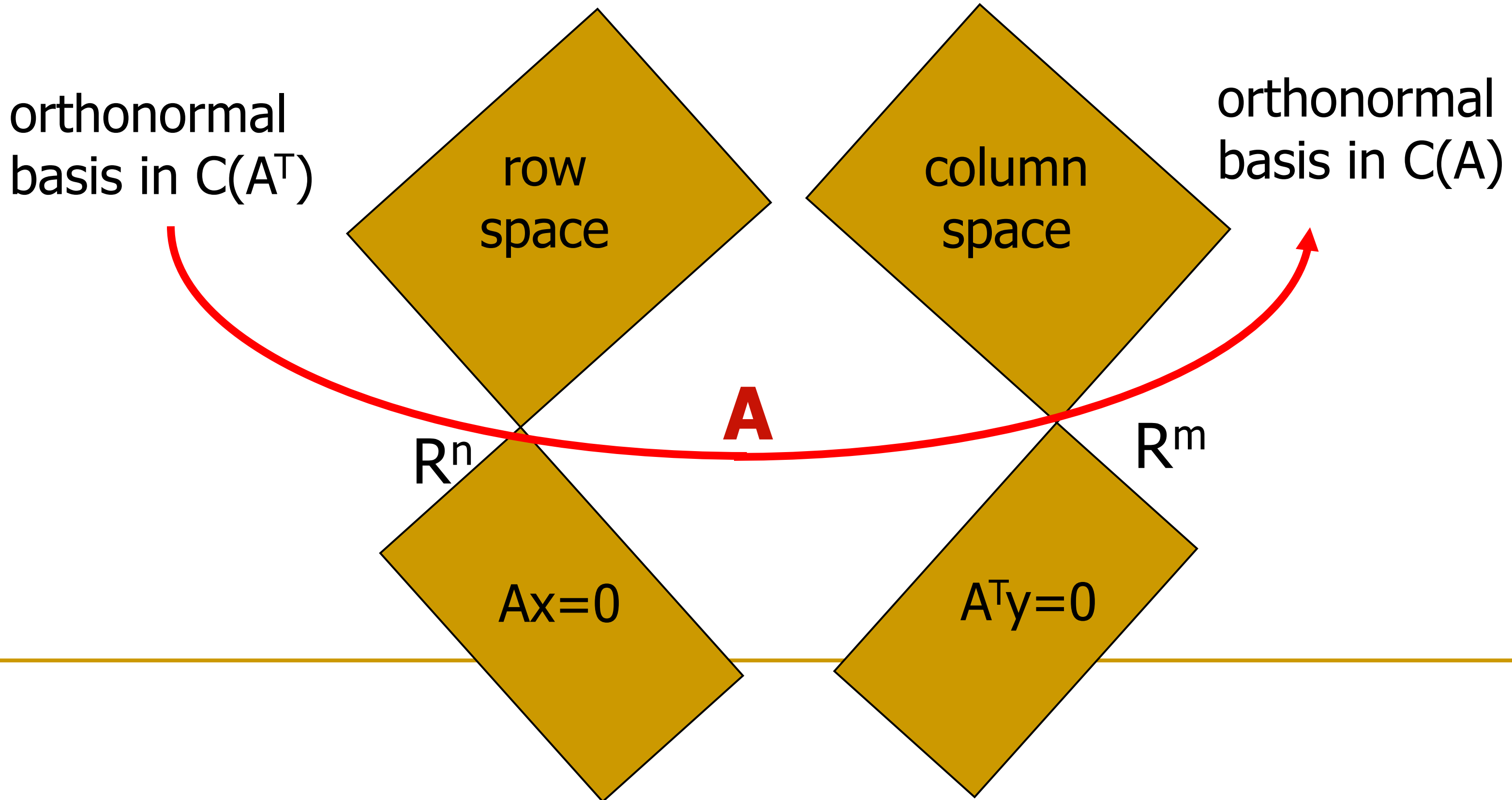
$A:$ any matrix
$A = U\Sigma V^T$
$A^{-1} = V\Sigma^{-1}U^T$
$x = V\Sigma^{-1}U^T b$

Clearly
the winner

Ideas Behind SVD

There are many choices of basis in $C(A^T)$ and $C(A)$, but we want the orthonormal ones

- Goal: for $A_{m \times n}$
  - find orthonormal bases for $C(A^T)$ and $C(A)$



SVD (2X2)

I haven't told you how to find v_i's (p.9)

Assume $r = 2$

unit vectors in $C(A^T): v_1 \perp v_2$

AND we want their images in $C(A): Av_1 \perp Av_2$

$$u_1 \equiv \frac{Av_1}{\|Av_1\|} = \frac{Av_1}{\sigma_1}, u_2 \equiv \frac{Av_2}{\|Av_2\|} = \frac{Av_2}{\sigma_2}$$

σ: represent the *length* of images; hence non-negative

$$A\begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 u_1 & \sigma_2 u_2 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix}$$

$$AV = U\Sigma$$

SVD 2x2 (cont)

$$AV = U\Sigma$$

$$A = U\Sigma V^{-1} = U\Sigma V^{T}$$

Another diagonalization using 2 sets of orthogonal bases

Compare

When A has complete set of e-vectors, we have

$AS = S\Lambda$, $A = S\Lambda S^{-1}$

but S in general is not orthogonal

When A is symmetric, we have

$A = Q\Lambda Q^{T}$

Why are orthonormal bases good?

- ( )$^{-1}$=( )^{T}
- Implication:
  - Matrix inversion
  - Ax=b

$$A = U\Sigma V^T$$

$$A^{-1} = \left(U\Sigma V^T\right)^{-1} = V\Sigma^{-1}U^T$$

$$A = U\Sigma V^T$$

$$Ax = b$$

$$\left(U\Sigma V^T\right)x = b$$

$$\Sigma\left(V^T x\right) = \left(U^T b\right) \ldots \text{diagonal system}$$

More on U and V

$$A^T A = V\Sigma^T U^T U \Sigma V^T = V \begin{bmatrix} \sigma_1^2 & \\ & \sigma_2^2 \end{bmatrix} V^T$$

V: eigenvectors of A^TA

U: eigenvectors of AA^T

Similarly,

$$AA^T = U\Sigma V^T V \Sigma^T U^T = U \begin{bmatrix} \sigma_1^2 & \\ & \sigma_2^2 \end{bmatrix} U^T$$

[Caution]: they (u_i and v_i) may differ in signs …
Find v_i first, then use Av_i to find u_i

This is the key to solve SVD

SVD: $A=U\Sigma V^T$

- The singular values are the diagonal entries of the Σ matrix and are arranged in *descending* order
- The singular values are always real (non-negative) numbers
- If A is real matrix, U and V are also real

Example (2x2, full rank)

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

$$A^T A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}, v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, v_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$Av_1 = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} = \sigma_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}, u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$Av_2 = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix} = \sigma_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

STEPS:

1. Find e-vectors of A^TA; normalize the basis
2. Compute Av_i, get σ_i
   If $\sigma_i \neq 0$, get u_i
   Else find u_i from $N(A^T)$

$$A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

SVD Theory

$$AV = U\Sigma$$

$$\rightarrow Av_j = \sigma_j u_j,\ j = 1,2,\cdots,r$$

- If σ_j=0, Av_j=0→v_j is in N(A)
  - The corresponding u_j in N(A^T)
    - [$U^TA=\Sigma V^T=0$]
- Else, v_j in C(A^T)
  - The corresponding u_j in C(A)
- #of nonzero σ_j = rank

Example (2x2, rank deficient)

$$A=\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}, r=1, C(A^T)\text{ basis}: \begin{bmatrix}1\\ 1\end{bmatrix}; v_1=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix}$$

$$Av_1=\sigma_1 u_1$$

$$\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix}=\sigma_1\frac{1}{\sqrt{5}}\begin{bmatrix}2\\ 1\end{bmatrix}\rightarrow\sigma_1=\sqrt{10}$$

Can also be obtained from e-vectors of $A^T A$

$$v_2\perp v_1\rightarrow v_2\in N(A)\rightarrow Av_2=0\rightarrow v_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix}$$

$$u_2\perp u_1\rightarrow u_2\in N(A^T)\rightarrow A^T u_2=0\rightarrow u_2^T A=0\rightarrow u_1=\frac{1}{\sqrt{5}}\begin{bmatrix}1\\ -2\end{bmatrix}$$

$$\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}=U\Sigma V^T=\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

Example (cont)

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = U\Sigma V^T = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$A = U\Sigma V^T = \begin{bmatrix} u_1 & u_2 \end{bmatrix}\begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$$

$$= \begin{bmatrix} u_1 & u_2 \end{bmatrix}\begin{bmatrix} \sigma_1 v_1^T \\ 0 \end{bmatrix} = \sigma_1 u_1 v_1^T$$

Problem: Sign and correspondence (if more than one u_i for N(A)) matter?

Bases of N(A) and N(A^T) (u_2 and v_2 here) do not contribute the final result. The are computed to make U and V orthogonal.

Extend to A_{mxn}

$$Av_i = \sigma_i u_i, i = 1, \cdots, r$$

$$A\begin{bmatrix} v_1 & \cdots & v_r & v_{r+1} \cdots v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r & u_{r+1} \cdots u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}$$

Basis of N(A) Basis of N(A^T)

Dimension Check

$$AV = U\Sigma$$

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V_{n\times n}^T$$

$$\left(A^T A\right)_{n\times n} = \left(U\Sigma V^T\right)^T\left(U\Sigma V^T\right) = V_{n\times n}\left(\Sigma^T\Sigma\right)_{n\times n}V_{n\times n}^T$$

$$\left(AA^T\right)_{m\times m} = \left(U\Sigma V^T\right)\left(U\Sigma V^T\right)^T = U_{m\times m}\left(\Sigma^T\Sigma\right)_{m\times m}V_{m\times m}^T$$

Extend to A_{mxn} (cont)

$$A\begin{bmatrix} v_1 & \cdots & v_r & v_{r+1}\cdots v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r & u_{r+1}\cdots u_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}$$

Summation of *r* rank-one matrices!

$$A = \begin{bmatrix} \sigma_1 u_1 & \cdots & \sigma_r u_r & 0\cdots 0 \end{bmatrix}\begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \\ v_{r+1}^T \\ \vdots \\ v_m^T \end{bmatrix} = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$$

Bases of N(A) and N(A^T) do not contribute

They are useful only for nullspace solutions

$$A=\begin{bmatrix}1&1&0\\0&1&1\end{bmatrix}, r=2, A_{2\times3}V_{3\times3}=U_{2\times2}\Sigma_{2\times3}$$

$$V: A^TA=\begin{bmatrix}1&1&0\\1&2&1\\0&1&1\end{bmatrix}$$

$$Av_1=\frac{1}{\sqrt{6}}\begin{bmatrix}3\\3\end{bmatrix}=\frac{3\sqrt{2}}{\sqrt{6}}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \sigma_1=\sqrt{3}$$

$$Av_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}, \sigma_2=1$$

$$\lambda_1=3, v_1=\frac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\1\end{bmatrix}$$

$$\lambda_2=1, v_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix}$$

$$\lambda_3=0, v_3=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix} \text{(nullspace)}$$

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

C(A^T) N(A) C(A)

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

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}, r = 2, A_{3\times2}V_{2\times2} = U_{3\times3}\Sigma_{3\times2}$$

$$V : A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$Av_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \frac{\sqrt{6}}{\sqrt{2}}\frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \sqrt{3}\frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \sigma_1 = \sqrt{3}$$

$$\lambda_1 = 3, v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$Av_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \sigma_2 = 1$$

It's alright if we fill N(A^T) with 0 if we only care about C(A^T) and C(A), the ones correspond to nonzero σ's

$$\lambda_2 = 1, v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$Av_3 = 0$ basis of $N(A^T)$...obtained in e-vector of AA^T, or ortherwise

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

C(A^T) C(A) N(A^T)

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

Summary

- SVD chooses the right basis for the 4 subspaces
- $AV=U\Sigma$
  - $v_1 \ldots v_r$: orthonormal basis in R^n for $C(A^T)$
  - $v_{r+1} \ldots v_n$: $N(A)$
  - $u_1 \ldots u_r$: in R^m $C(A)$
  - $u_{r+1} \ldots u_m$: $N(A^T)$
- These bases are not only $\perp$, but also $Av_i = \sigma_i u_i$
- High points of Linear Algebra
  - Dimension, rank, orthogonality, basis, diagonalization, ...

SVD Applications

- Using SVD in computation, rather than A, has the advantage of being more robust to numerical error
- Many applications:
  - Image compression
  - Solve Ax=b for all cases (unique, many, no solutions; least square solutions)
  - rank determination, matrix approximation, …
- SVD usually found by iterative methods (see Numerical Recipe, Chap.2)

SVD and Ax=b (m≥n)

- Check for existence of solution

$$U\Sigma V^T x = b \rightarrow \Sigma \underbrace{V^T x}_{z} = \underbrace{U^T b}_{d}$$

$$\Sigma z = d$$

If $\sigma_i = 0$ but $d_i \neq 0$,

solution does not exist

Ax=b (inconsistent)

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = U\Sigma V^T = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = U^T b = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 8 \\ 3 \end{bmatrix} = \begin{bmatrix} 20/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

No solution!

Ax=b (underdetermined)

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

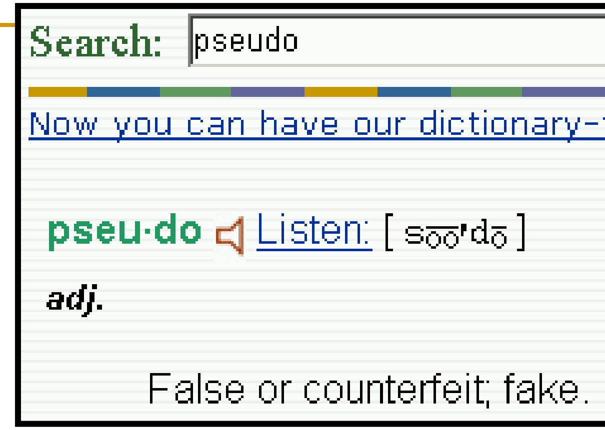
$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = U\Sigma V^T = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = U^T b = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 20/\sqrt{5} \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} \rightarrow x_{particular} = Vz = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$x_{complete} = x_{particular} + x_{null} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + c \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Pseudo Inverse (Sec7.4, p.395)



- The role of A:
  - Takes a vector v_i from row space to $\sigma_i u_i$ in the column space
- The role of A^{-1} (if it exists):
  - Does the opposite: takes a vector u_i from column space to row space v_i

$$Av_i = \sigma_i u_i$$
$$v_i = \sigma_i A^{-1} u_i$$
$$A^{-1} u_i = \frac{1}{\sigma_i} v_i$$

Fig 7.4, p.396

Pseudo Inverse (cont)

- While A^{-1} may not exist, a matrix that takes u_i back to v_i/σ_i does exist. It is denoted as A^+, the pseudo inverse
- A^+: dimension n by m

$$A^+ u_i = \frac{1}{\sigma_i} v_i \quad \text{for } i \le r \quad \text{and} \quad A^+ u_i = 0 \ \text{ for } i > r$$

$$A^+ = V_{n\times n} \Sigma^+_{n\times m} U^T_{m\times m}$$

$$= \begin{bmatrix} v_1 \cdots & v_r & \cdots v_n \end{bmatrix} \begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \\ & & \end{bmatrix} \begin{bmatrix} u_1 \cdots & u_r & \cdots u_m \end{bmatrix}^T$$

Pseudo Inverse and Ax=b

$$Ax = b$$

$$x = A^{+}b = V\Sigma^{+}U^{T}b$$

A panacea for Ax=b

- Full rank: A^{-1} exist; A^{+} is the same as A^{-1}
- Underdetermined case: many solutions, but will find the one with the smallest magnitude |x|
- Overdetermined case: find the solution that minimize the error r=|Ax–b|, the least square solution
- [proofs given below]

[Proofs from NR-1]

Proof: Consider $|\mathbf{x} + \mathbf{x}'|$, where $\mathbf{x}'$ lies in the nullspace. Then, if $\mathbf{W}^{-1}$ denotes the modified inverse of $\mathbf{W}$ with some elements zeroed,

$$\begin{aligned} |\mathbf{x} + \mathbf{x}'| &= \left|\mathbf{V} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}^T \cdot \mathbf{b} + \mathbf{x}'\right| \\ &= \left|\mathbf{V} \cdot (\mathbf{W}^{-1} \cdot \mathbf{U}^T \cdot \mathbf{b} + \mathbf{V}^T \cdot \mathbf{x}')\right| \\ &= \left|\mathbf{W}^{-1} \cdot \mathbf{U}^T \cdot \mathbf{b} + \mathbf{V}^T \cdot \mathbf{x}'\right| \end{aligned} \tag{2.6.8}$$

Here the first equality follows from (2.6.7), the second and third from the orthonormality of $\mathbf{V}$. If you now examine the two terms that make up the sum on the right-hand side, you will see that the first one has nonzero j components only where $w_j \neq 0$, while the second one, since $\mathbf{x}'$ is in the nullspace, has nonzero j components only where $w_j = 0$. Therefore the minimum length obtains for $\mathbf{x}' = 0$, q.e.d.

[Proofs from NR-2]

The proof is similar to (2.6.8): Suppose we modify $\mathbf{x}$ by adding some arbitrary $\mathbf{x}'$. Then $\mathbf{A} \cdot \mathbf{x} - \mathbf{b}$ is modified by adding some $\mathbf{b}' \equiv \mathbf{A} \cdot \mathbf{x}'$. Obviously $\mathbf{b}'$ is in the range of $\mathbf{A}$. We then have

$$\begin{aligned}
\left|\mathbf{A} \cdot \mathbf{x} - \mathbf{b} + \mathbf{b}'\right| &= \left|(\mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T) \cdot (\mathbf{V} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}^T \cdot \mathbf{b}) - \mathbf{b} + \mathbf{b}'\right| \\
&= \left|(\mathbf{U} \cdot \mathbf{W} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}^T - \mathbf{1}) \cdot \mathbf{b} + \mathbf{b}'\right| \\
&= \left|\mathbf{U} \cdot \left[(\mathbf{W} \cdot \mathbf{W}^{-1} - \mathbf{1}) \cdot \mathbf{U}^T \cdot \mathbf{b} + \mathbf{U}^T \cdot \mathbf{b}'\right]\right| \\
&= \left|(\mathbf{W} \cdot \mathbf{W}^{-1} - \mathbf{1}) \cdot \mathbf{U}^T \cdot \mathbf{b} + \mathbf{U}^T \cdot \mathbf{b}'\right|
\end{aligned} \tag{2.6.10}$$

Now, $(\mathbf{W} \cdot \mathbf{W}^{-1} - \mathbf{1})$ is a diagonal matrix which has nonzero j components only for $w_j = 0$, while $\mathbf{U}^T \mathbf{b}'$ has nonzero j components only for $w_j \neq 0$, since $\mathbf{b}'$ lies in the range of $\mathbf{A}$. Therefore the minimum obtains for $\mathbf{b}' = 0$, q.e.d.

Ex: full rank

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}, b = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

$$A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$Ax = b \rightarrow U\Sigma V^T x = b \rightarrow x = V\mathrm{diag}(1/\sigma_i)U^T b$$

$$x = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Ex: under-determined

$$Ax = b \qquad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

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

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

$$x = A^{+}b = V\Sigma^{+}U^{T}b$$

$$= \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & x \\ \frac{2}{\sqrt{6}} & 0 & y \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & z \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \\ & 1 \\ & \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ \frac{2}{3} \\ \frac{-2}{3} \end{bmatrix}$$

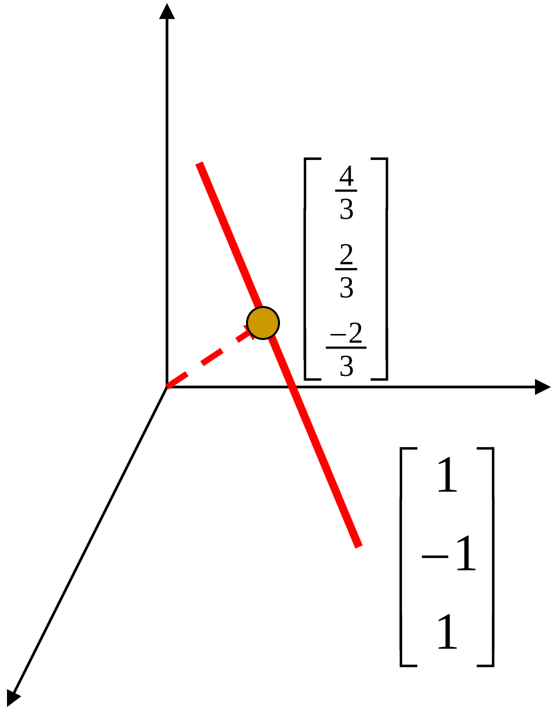
Under-determined (cont)

$$\begin{cases} x + y \quad\quad = 2 \\ \quad\quad y + z = 0 \end{cases}$$

Complete Solution :

$$x = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ Geometrically, a line in } R^3$$

$$x = A^{+} b \text{ gives a particular with smallest } \|x\|$$



Ex: over-determined

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, r = 2, A_{3\times2}V_{2\times2} = U_{3\times3}\Sigma_{3\times2}$$

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

$$A = U\Sigma V^T$$

$$A^+ = V\Sigma^+ U^T$$

Will show this need not be computed…

Over-determined (cont)

$$x = V\Sigma^{+}U^{T}b$$

$$= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{3}} & & \\ & 1 & \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ x & y & z \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{3}} & & \\ & 1 & \end{bmatrix}\begin{bmatrix} \frac{4}{\sqrt{6}} \\ \frac{2}{\sqrt{2}} \\ x+2y-z \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \frac{2\sqrt{2}}{3} \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{-1}{3} \end{bmatrix}$$

Compare $A^{T}A\hat{x} = A^{T}b$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{-1}{3} \end{bmatrix}$$

Same result!!

Ex: general case, no solution

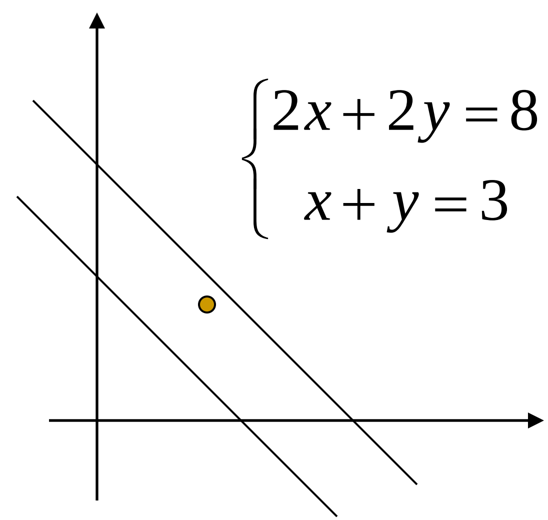
$$A=\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}, b=\begin{bmatrix}8\\ 3\end{bmatrix}$$

$$\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}=U\Sigma V^T=\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

$$x=V\Sigma^+U^Tb$$

$$=\begin{bmatrix}\frac{1}{\sqrt{2}} & x\\ \frac{1}{\sqrt{2}} & y\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{10}} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\\ x & y\end{bmatrix}\begin{bmatrix}8\\ 3\end{bmatrix}$$

$$=\begin{bmatrix}\frac{1}{2\sqrt{5}} & 0\\ \frac{1}{2\sqrt{5}} & 0\end{bmatrix}\begin{bmatrix}\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\\ x & y\end{bmatrix}\begin{bmatrix}8\\ 3\end{bmatrix}=\begin{bmatrix}\frac{1}{5} & \frac{1}{10}\\ \frac{1}{5} & \frac{1}{10}\end{bmatrix}\begin{bmatrix}8\\ 3\end{bmatrix}=\begin{bmatrix}\frac{19}{10}\\ \frac{19}{10}\end{bmatrix}$$



$$\begin{cases}2x+2y=8\\ x+y=3\end{cases}$$

Matrix Approximation

$A_i = U\Sigma_i V^T$

Σ_i : the rank i version of Σ (by setting last $m-i$ σ's to zero)

A_i : the best rank i approximation to A in the sence of Euclidean distance

$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_m u_m v_m^T$

Storage save : rank one matrix $(m+n)$ numbers

Operation save : $m+n$

making small σ's to zero and back substitute
(see next page for application in image compression)

Image Compression

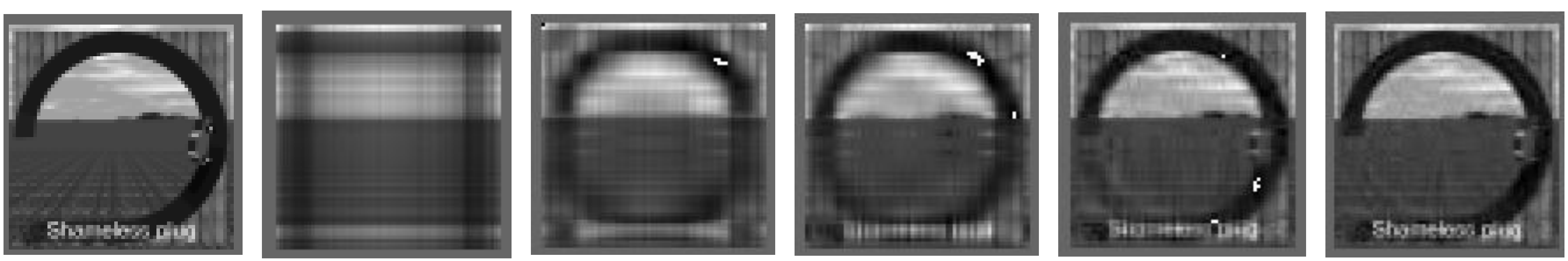
Play with the demo program

- As described in text p.352
- For grey scale images: m×n bytes

After SVD, taking the most significant r terms:

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

- Only need to store $r\times$(m+n+1)



Original

64×64

r = 1,3,5,10,16 (no perceivable difference afterwards)