

THE SINGULAR VALUE DECOMPOSITION

- The SVD – existence - properties.
- Pseudo-inverses and the SVD
- Use of SVD for least-squares problems
- Applications of the SVD
- Text: mainly sect. 2.4

The Singular Value Decomposition (SVD)

Theorem For any matrix $A \in \mathbb{R}^{m \times n}$ there exist unitary matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that

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

where Σ is a diagonal matrix with entries $\sigma_{ii} \geq 0$.

$$\sigma_{11} \geq \sigma_{22} \geq \cdots \sigma_{pp} \geq 0 \text{ with } p = \min(n, m)$$

➤ The σ_{ii} are the **singular values** of A .

➤ σ_{ii} is denoted simply by σ_i

Proof: Let $\sigma_1 = \|A\|_2 = \max_{x, \|x\|_2=1} \|Ax\|_2$. There exists a pair of unit vectors v_1, u_1 such that

$$Av_1 = \sigma_1 u_1$$

8-2

Csci 5304 – October 29, 2013

➤ Complete v_1 into an orthonormal basis of \mathbb{R}^n

$$V \equiv [v_1, V_2] = n \times n \text{ unitary}$$

➤ Complete u_1 into an orthonormal basis of \mathbb{R}^m

$$U \equiv [u_1, U_2] = m \times m \text{ unitary}$$

➤ Then, it is easy to show that

$$AV = U \times \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \rightarrow U^T AV = \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \equiv A_1$$

➤ Observe that

$$\left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \geq \sigma_1^2 + \|w\|^2 = \sqrt{\sigma_1^2 + \|w\|^2} \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2$$

➤ This shows that w must be zero [why?]

➤ Complete the proof by an induction argument.

Case 1:

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

Case 2:

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

8-3

Csci 5304 – October 29, 2013

8-4

Csci 5304 – October 29, 2013

The “thin” SVD

- Consider the Case-1. It can be rewritten as

$$A = [U_1 U_2] \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} V^T$$

Which gives:

$$A = U_1 \Sigma_1 V^T$$

where U_1 is $m \times n$ (same shape as A), and Σ_1 and V are $n \times n$

- referred to as the “thin” SVD. Important in practice.

📌 How can you obtain the thin SVD from the QR factorization of A and the SVD of an $n \times n$ matrix?

8-5

Csci 5304 – October 29, 2013

A few properties. Assume that

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 \text{ and } \sigma_{r+1} = \dots = \sigma_p = 0$$

Then:

- $\text{rank}(A) = r = \text{number of nonzero singular values.}$
- $\text{Ran}(A) = \text{span}\{u_1, u_2, \dots, u_r\}$
- $\text{Null}(A) = \text{span}\{v_{r+1}, v_{r+2}, \dots, v_n\}$
- The matrix A admits the SVD expansion:

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T$$

8-6

Csci 5304 – October 29, 2013

Properties of the SVD (continued)

- $\|A\|_2 = \sigma_1 = \text{largest singular value}$
- $\|A\|_F = (\sum_{i=1}^r \sigma_i^2)^{1/2}$
- When A is an $n \times n$ nonsingular matrix then $\|A^{-1}\|_2 = 1/\sigma_n = \text{inverse of smallest s.v.}$

Let $k < r$ and

$$A_k = \sum_{i=1}^k \sigma_i u_i v_i^T$$

then

$$\min_{\text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

8-7

Csci 5304 – October 29, 2013

Right and Left Singular vectors:

$$\begin{aligned} A v_i &= \sigma_i u_i \\ A^T u_j &= \sigma_j v_j \end{aligned}$$

- Consequence $A^T A v_i = \sigma_i^2 v_i$ and $A A^T u_i = \sigma_i^2 u_i$
- Right singular vectors (v_i 's) are eigenvectors of $A^T A$
- Left singular vectors (u_i 's) are eigenvectors of $A A^T$
- Possible to get the SVD from eigenvectors of $A A^T$ and $A^T A$ – but: difficulties due to non-uniqueness of the SVD

8-8

Csci 5304 – October 29, 2013

Define the $r \times r$ matrix

$$\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_r)$$

► Let $A \in \mathbb{R}^{m \times n}$ and consider $A^T A \in \mathbb{R}^{n \times n}$:

$$A^T A = V \Sigma^T \Sigma V^T \rightarrow A^T A = V \underbrace{\begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}}_{n \times n} V^T$$

► This gives the spectral decomposition of $A^T A$.

► Similarly, U gives the eigenvectors of AA^T .

$$AA^T = U \underbrace{\begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}}_{m \times m} U^T$$

Important:

$A^T A = V D_1 V^T$ and $AA^T = U D_2 U^T$ give the SVD factors U, V up to signs!

Pseudo-inverse of an arbitrary matrix

The pseudo-inverse of A is given by

$$A^\dagger = V \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^T = \sum_{i=1}^r \frac{v_i u_i^T}{\sigma_i}$$

Moore-Penrose conditions:

The pseudo inverse of a matrix is uniquely determined by these four conditions:

- (1) $AXA = A$
- (2) $XAX = X$
- (3) $(AX)^H = AX$
- (4) $(XA)^H = XA$

► In the full-rank overdetermined case, $A^\dagger = (A^T A)^{-1} A^T$

Least-squares problems and the SVD

► SVD can give much information about solving overdetermined and underdetermined linear systems.

Let A be an $m \times n$ matrix and $A = U \Sigma V^T$ its SVD with $r = \text{rank}(A)$, $V = [v_1, \dots, v_n]$ $U = [u_1, \dots, u_m]$. Then

$$x_{LS} = \sum_{i=1}^r \frac{u_i^T b}{\sigma_i} v_i$$

minimizes $\|b - Ax\|_2$ and has the smallest 2-norm among all possible minimizers. In addition,

$$\rho_{LS} \equiv \|b - Ax_{LS}\|_2 = \|z\|_2 \text{ with } z = [u_{r+1}, \dots, u_m]^T b$$

Least-squares problems and pseudo-inverses

- A restatement of the first part of the previous result:

Consider the general linear least-squares problem

$$\min_{x \in S} \|x\|_2, \quad S = \{x \in \mathbb{R}^n \mid \|b - Ax\|_2 \text{ min}\}.$$

This problem always has a unique solution given by

$$x = A^\dagger b$$

8-13

Csci 5304 – October 29, 2013

✎ Consider the matrix:

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

- Compute the singular value decomposition of A
- Find the matrix B of rank 1 which is the closest to the above matrix in the 2-norm sense.
- What is the pseudo-inverse of A ?
- What is the pseudo-inverse of B ?
- Find the vector x of smallest norm which minimizes $\|b - Ax\|_2$ with $b = (1, 1)^T$
- Find the vector x of smallest norm which minimizes $\|b - Bx\|_2$ with $b = (1, 1)^T$

8-14

Csci 5304 – October 29, 2013

Ill-conditioned systems and the SVD

- Let A be $m \times m$ and $A = U\Sigma V^T$ its SVD
- Solution of $Ax = b$ is $x = A^{-1}b = \sum_{i=1}^m \frac{u_i^T b}{\sigma_i} v_i$
- When A is very ill-conditioned, it has many small singular values. The division by these small σ_i 's will amplify any noise in the data. If $\tilde{b} = b + \epsilon$ then

$$A^{-1}\tilde{b} = \sum_{i=1}^m \frac{u_i^T b}{\sigma_i} v_i + \underbrace{\sum_{i=1}^m \frac{u_i^T \epsilon}{\sigma_i} v_i}_{\text{Error}}$$

- Result: solution could be completely meaningless.

8-15

Csci 5304 – October 29, 2013

Remedy: SVD regularization

Truncate the SVD by only keeping the σ_i 's that $\geq \tau$, where τ is a threshold

- Gives the Truncated SVD solution (**TSVD solution**):

$$x_{TSVD} = \sum_{\sigma_i \geq \tau} \frac{u_i^T b}{\sigma_i} v_i$$

- Many applications [e.g., Image processing,..]

8-16

Csci 5304 – October 29, 2013

Numerical rank and the SVD

- Assume that the original matrix A is exactly of rank k .
- The **computed** SVD of A will be the SVD of a nearby matrix $A + E$.
- Easy to show that $|\hat{\sigma}_i - \sigma_i| \leq \alpha \sigma_1 \underline{u}$
- Result: zero singular values will yield small computed singular values
- Determining the “numerical rank:” treat singular values below a certain threshold δ as zero.
- Practical problem : need to set δ .