Lecture IX
Singular-Value Decomposition

- **SVD: notion, examples and properties**

- **Some Useful Matrix Inequalities**

Goal of SVD

**Decomposition of a, not necessarily square, matrix into a product of three matrices:
one containing its singular values, two unitary.**

Note 1: Extension of canonical forms for square matrices
Note 2: Used to define a pseudo-inverse for a nonsquare or a singular matrix

The Fundamental Theorem

Consider $A \in \mathbb{C}^{m\times n}$, with rank$A = r$. Then, there exist unitary matrices $V \in \mathbb{C}^{m\times m}$, $W \in \mathbb{C}^{n\times n}$ such that

$$A = V\Sigma W^*$$

where

$$\Sigma = \begin{pmatrix} S_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix},$$

$$S = diag(\sigma_i), \quad \sigma_i = \sqrt{\lambda_i(A^* A)},$$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, singular values of A.

Comments

♥ SVD is not unique!

♣ Let $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$, $W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}$.

$$A = V\Sigma W^*$$

$\Rightarrow$

$$A = V_1 S W_1^* \quad \text{(Reduced SVD)}$$

Construction of *V* and *W*

(1) First, decompose the $n \times n$ (Hermitian and positive semifinite) matrix $A^* A$ using its orthonormal set of eigenvectors

$$\tilde{W}_1 = \begin{bmatrix} w_1 & \cdots & w_r \end{bmatrix}, \ \tilde{W}_2 = \begin{bmatrix} w_{r+1} & \cdots & w_n \end{bmatrix}$$

i.e. $\tilde{W}^* A^* A \tilde{W} = \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}.$ (CDF)

Then, $\tilde{W}_1^* A^* A \tilde{W}_1 = S^2 \Rightarrow S^{-1} \tilde{W}_1^* A^* A \tilde{W}_1 S^{-1} = I.$

From (CDF), it follows that $\tilde{W}_2^* A^* A \tilde{W}_2 = 0 \Rightarrow A \tilde{W}_2 = 0.$

Define $V_1 \triangleq A \tilde{W}_1 S^{-1}$, implying $V_1 V_1^* = I.$

Construction of *V* and *W*

(2) Now, choose <u>any</u> $V_2 \in \mathbb{C}^{m\times(m-r)}$ such that

$[V_1 \ \ V_2]$ is unitary.

Then, it can be directly checked

that $V^* A W = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}$, with $W = \tilde{W}$.

Examples

Can you find SVD for the following matrices

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}?$$

Answer

$$A_1 = UIU^*$$

with $U \in \mathbb{C}^{2\times 2}$ any arbitrary unitary matrix.

$$A_2 = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Exercise

Verify that
$$\begin{pmatrix} \frac{1}{3} & \frac{-2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{-\sqrt{5}}{3} \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

is an SVD of $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 2 & 2 \end{pmatrix}$.

Case of Symmetric Matrices

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix.

Question : *What* is an SVD of A?

Answer

A can be made diagonal by means of an orthogonal matrix O composed of orthonormal eigenvectors, i.e., $O^T AO = \Lambda = diag\left(\lambda_i\right)$.

Therefore, $O\Lambda O^T$ is an SVD of A.

Properties

Assume SVD $A = V\Sigma W^*$. Then,

(1) rank(A) = the no. of nonzero singular values of A.

(2) A has the dyadic (or outer product) expansion:

$$A = \sum_{i=1}^{r} \sigma_i v_i w_i^*$$

(3) The singular vectors satisfy

$$\begin{cases} Aw_i = \sigma_i v_i, \\ A^* v_i = \sigma_i w_i. \end{cases}$$

Moore-Penrose Generalized Inverse

Question :

How can we define a (generalized) inverse, denoted $A^{\dagger}$, of a matrix A which may be singular or nonsquare?

Consider $A:\mathbb{C}^n \to \mathbb{C}^m$, denoted $A \in \mathbb{C}^{m\times n}$.

Define $T : N(A)^{\perp} \to R(A)$ by

$$Tx = Ax, \text{ for all } x \in N(A)^{\perp}.$$

Moore-Penrose Generalized Inverse

Consider $A:\mathbb{C}^n \to \mathbb{C}^m$, denoted $A \in \mathbb{C}^{m\times n}$.

Define $T: N(A)^{\perp} \to R(A)$ by

$$Tx = Ax, \text{ for all } x \in N(A)^{\perp}.$$

Definition of the Moore-Penrose Pseudoinverse

$A^{+}: \mathbb{C}^m \to \mathbb{C}^n$

$$A^{+}y = T^{-1}y_1,$$

with $y = y_1 + y_2,\ y_1 \in R(A),\ y_2 \in R(A)^{\perp}$.

Moore-Penrose Generalized Inverse: Computation

Consider $A \in \mathbb{C}^{m \times n}$ and its SVD $A := V \sum W^*$.
Then, $A^\dagger := W \sum^\dagger V^*$, with $\sum^+$ being the transpose of $\sum$ in which the positive singular values of A are replaced by their reciprocals.

Properties of Moore-Penrose Generalized Inverse

(1) AA^+ and A^+A are Hermitian.

(2) $AA^+A = A$.

(3) $A^\dagger AA^\dagger = A^\dagger$.

(4) $A^\dagger = A^{-1}$, if A is square and nonsingular.

(5) $A^\dagger$ always exists and is the unique matrix that satisfies the same properties $(1)-(3)$.

See: R. Penrose, "A generalized inverse for matrices", 1955.

Proof on Uniqueness

Assume that $X \in \mathbb{C}^{n\times m}$ is a matrix satisfying the above properties (1)-(3), i.e. AX and XA are Hermitian, $AXA = A$, and $XAX = X$. *Then,*

$$X = XAX = X\left(AX\right)^{*} = XX^{*}A^{*} = XX^{*}\left(AA^{+}A\right)^{*} = XX^{*}A^{*}A^{+*}A^{*}$$

$$= X\left(AX\right)^{*}\left(AA^{+}\right)^{*} = XAXAA^{\dagger} = XAA^{\dagger} = \left(XA\right)^{*}A^{\dagger} = A^{*}X^{*}A^{?}$$

$$= \left(AA^{\dagger}A\right)^{*}X^{*}A^{+} = A^{*}A^{+*}A^{*}X^{*}A^{+} = \left(A^{?}A\right)^{*}\left(XA\right)^{*}A^{+}$$

$$= A^{\dagger}AXAA^{\dagger} = A^{\dagger}AA^{\dagger} = A^{\dagger},$$ END OF PROOF

Numerical Result

$$A^{+} = \lim_{\delta\to 0}\left(A^{T}A+\delta^{2}I\right)^{-1}A^{T} = \lim_{\delta\to 0}A^{T}\left(A^{T}A+\delta^{2}I\right)^{-1}$$

See: A. Albert, Regression and the Moore-Penrose Pseudoinverse, p.19, 1972.

Examples

What is the Moore-Penrose Pseudoinverse of

1) a scalar $a \in \mathbb{R}$?

2) a vector $v \in \mathbb{R}^n$?

3) a matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$?

Answers

1) For any scalar a, $a^+ = a^{-1}$ if $a \neq 0$, $a^+ = 0$ if $a = 0$.

2) For any vector $v \in \mathbb{R}^n$,

$$v^+ = \left(v^T v\right)^+ v^T = v^T / v^T v \text{ if } v \neq 0, \ = 0 \text{ if not.}$$

3) $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^+ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$

Application to Linear Equations

Consider linear equations

$Ax = b,$ with $A \in \mathbb{C}^{m\times n},\ b \in \mathbb{C}^{m}, rank(A) = k.$

By means of SVD

$$A = V\sum W^{*} \Rightarrow \sum(W^{*}x) = V^{*}b.$$

Then the last $m-k$ rows of $\sum$ are 0, and hence it is necessary and sufficient for existence of solutions that the last $m-k$ entries of $V^{*}b$ are zero.

In other words, b is orthogonal to the last $m-k$ left singular vectors of A.

$(Cont'd)$

Then, letting $V = [v_1 \cdots v_m]$, $W = [w_1 \cdots w_m]$,

$$\Sigma(W^* x) = V^* b$$

$$\Rightarrow \quad (W^* x)^* = \left[\frac{b^* v_1}{\sigma_1}, \cdots, \frac{b^* v_k}{\sigma_k}, 0, \cdots, 0 \right]^*$$

$$\Rightarrow \; x = \sum_{i=1}^{k} \frac{v_i^* b}{\sigma_i} w_i \text{ is a solution to } Ax = b.$$

The general solutions of $Ax = 0$ take the form

$$x_h = \sum_{i=k+1}^{n} c_i w_i \quad \text{(do you know why?)}$$

So, all solutions of $Ax = b$ are in the form

$$x = \sum_{i=1}^{k} \frac{v_i^* b}{\sigma_i} w_i + \sum_{i=k+1}^{n} c_i w_i \quad \text{(do you know why?)}$$

END.

Example -- Revisited

The following inhomogeneous equation

$$
\begin{aligned}
x_1 + 2x_2 &= 5 \\
2x_1 + 4x_2 &= 10 \\
3x_1 + 6x_2 &= 15
\end{aligned}
$$

has an infinite number of solutions

$$
\begin{aligned}
x &= x_p + x_h \\
&= \begin{bmatrix} 5 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \qquad \lambda \in \mathbb{R}.
\end{aligned}
$$

Matrix Inequalities

Matrix inequalities based on inner product play a vital role in systems science and engineering.

Cauchy-Schwarz Inequality

For any pair of vectors $x,\ y \in \mathbb{C}^n$,

$$\left|\langle x, y\rangle\right| \ \leq \ \|x\| \bullet \|y\|$$

where the equality holds if and only if $x,\ y$ are linearly dependent, i.e. $x = cy$ for some scalar c.

Sketch of Proof

It is equivalent to proving the following:

$$\left|\langle x, y\rangle\right| \doteq \left|\sum_{i=1}^{n} x_i \bar{y}_i\right| \le 1$$

for any unit vectors $x,\ y,$ with equality iff $x = cy$.

To prove the latter, notice that

$$\left|\sum x_i \bar{y}_i\right| \le \sum |x_i| \bullet |y_i|,$$ with equality iff $\arg\left(x_i \bar{y}_i\right) = \theta$ (angle, independent of i).

Sketch of Proof (cont'd)

So, $\left|\sum x_i \bar{y}_i\right| \le \sum |x_i| \bullet |y_i|$, with "=" iff $\arg\left(x_i \bar{y}_i\right) = \theta$

Using $|x_i| \bullet |y_i| \le \frac{1}{2}|x_i|^2 + \frac{1}{2}|y_i|^2$ with "=" iff $|x_i| = |y_i|$,

it follows that $\left|\sum x_i \bar{y}_i\right| \le \frac{1}{2} + \frac{1}{2} = 1$, as wished.

Moreover, the equality holds with:

$$x_i = e^{j\theta_i},\ y_i = e^{-j\theta} e^{j\theta_i}, \forall i = 1, \ldots, n$$

$\Leftrightarrow$ $\quad x = cy$, with $c = e^{-j\theta}$.

Another Simpler Proof

The claim is obvious for the special cases: $x = 0,\ or\ y = 0.$

For $x \neq 0,\ y \neq 0,$ for any real number λ, the following holds:

$$\left| x + \lambda y \right|^2 = \lambda^2 \left| y \right|^2 + 2\lambda \left\langle x, y \right\rangle + \left| x \right|^2 .$$

It has a unique solution or no solution if

$$4\left\langle x, y \right\rangle^2 - 4\left| x \right|^2 \left| y \right|^2 \leq 0.$$

The equality holds iff it has a unique solution, i.e., $x = -\lambda y.$

Hadamard's Inequality

- $\left|\det A\right| \le \left\|a^1\right\| \bullet \left\|a^2\right\| \bullet \cdots \bullet \left\|a^n\right\|$

  for any matrix $A = \left(a^1 \ a^2 \ \cdots \ a^n\right) \in \mathbb{C}^{n\times n}$.
- The equality holds if and only if

  $a^j = 0 \ \exists$ some j, *or* $\left\langle a^j, a^k \right\rangle = 0 \ \ \forall j \ne k$.

Example

$$\left|\det\begin{pmatrix} 1+i & 2-i \\ 3 & 4 \end{pmatrix}\right| = \left|-2+7i\right| = \sqrt{53}$$

$$< \sqrt{\left|1+i\right|^2 + 3^2}\sqrt{\left|2-i\right|^2 + 4^2} = \sqrt{231}.$$

Proof of Hadamard's Inequality

Without loss of generality, assume that

$$\alpha_j = \left\| a^j \right\| > 0, \quad \forall j = 1, 2, \ldots, n.$$

Also, assume that $B = \begin{bmatrix} b^1 & \cdots & b^n \end{bmatrix}$ is a solution

to the constrained optimization problem:

$$\max_{\tilde{B}} \left| \det \tilde{B} \right|, \text{ subject to } \left| \tilde{b}^j \right| = \alpha_j, \forall j = 1, \ldots, n.$$

Clearly, $\left| \det B \right| \geq \prod \alpha_i > 0$, because $\tilde{B} = diag\left(\alpha_i\right)$

is a candidate solution.

Proof (cont'd)

Expanding $\det B$ by column j yields:

$$\det B = b_{1j}c_{1j} + \cdots + b_{nj}c_{nj}$$

with $C = \left(c_{ij}\right)$ the cofactor matrix of B,

i.e., $c_{ij} = (-1)^{i+j} \det B_{ij}, \ \forall i, j.$

By Cauchy-Schwarz inequality,

$$|\det B| \le \left\| b^j \right\| \bullet \left\| y \right\| \doteq \alpha_j \left\| y \right\|, \quad y \doteq col\left(\overline{c}_{1j}, \cdots, \overline{c}_{nj}\right)$$

Also, $b^j = \mu_j y$ for scalar μ_j because maximizer B *guarantees* the equality.

Proof (cont'd)

From the fact $b^j = \mu_j y$ for scalar μ_j, we can conclude that the columns of B are orthogonal.

Indeed, $\forall k \neq j$,

$$\left\langle b^k, b^j \right\rangle = \left\langle b^k, \mu_j \bar{c}_j \right\rangle = \bar{\mu}_j \sum_{i=1}^{n} b_{ik} c_{ij} = \bar{\mu}_j \det \breve{B}$$

where $\breve{B}$ is the matrix by replacing the column j of B by a duplicate column k of B. Thus

$$\left\langle b^k, b^j \right\rangle = 0, \quad \forall k \neq j.$$

Proof (cont'd)

From the fact $\langle b^k, b^j \rangle = 0, \quad \forall k \neq j,$ we have

$$|\det B|^2 = \det B^* \det B = \det\left(B^* B\right)$$

$$= \det\left(\left\langle b^j, b^i \right\rangle\right) = \det\left(diag\left(\alpha_i^2\right)\right) = \prod_{j=1}^{n} \alpha_j^2$$

Of course,

$$|\det A| \leq |\det B| = \prod_{j=1}^{n} \alpha_j, \text{ as wished.}$$

The equality holds only when A is another maximal solution, so its columns are mutually orthogonal.

Exercises

1. Determine the SVDs of the matrices

$$A_1 = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

2. Let $A = \left(a_{ij}\right)_{n\times n}$ and $\rho_i > 0,\ 1 \le i \le n.$

Show that $\left|\det A\right|^2 \le \prod_{j=1}^{n} \rho_j^{-2} \left(\sum_{i=1}^{n} \rho_i^2 \left|a_{ij}\right|^2 \right)$

When does the equality hold?