Matrix Factorizations

Kim, Hyun-Min

Department of Mathematics, Pusan National University
E-mail:hyunmin@pusan.ac.kr
Phone: 510-1060, 2596, 010-3833-8200

Schedule for Lecture

- Introduction to Numerical Linear Algebra
  - What is the numerical linear algebra?
  - Why?
  - How?
- Matrix Factorizations
  - What are matrix factorizations?
  - Existence and uniqueness of matrix factorizations
  - Why?
  - How?

- Two matrices
  - LU, QR Factorizations
- Three matrices
  - Diagonalization
  - Jordan canonical form
  - Schur decompositions
  - Singular decomposition

Linear Algebra

Math.

Linear Algebra by S. Friedberg, A. Insel and L. Spence
Chapters 1, 2, 3, 4, 5, 6, 7

Linear Algebra

Math.

Linear Algebra by S. Friedberg, A. Insel and L. Spence
Chapters 1, 2, 3, 4, 5, 6, 7

Eng.

Elementary Linear Algebra by H. Anton and C. Rorres
Chapters 1, 2, 3, 4, 5, 6, 7, 8

Mathematical Programming

Math.

Mathematica, Maple, MATLAB

Mathematical Programming

Math.

Mathematica, Maple, MATLAB

Eng.

Fortran, C, MATLAB

Errors

Definition

x: the true value

x^*: an approximation to x

- $|x - x^*|$: Absolute Error
- $\dfrac{|x - x^*|}{|x|}$ $(x \neq 0)$: Relative Error

Errors

Definition

x: the true value

x^*: an approximation to x

- $|x - x^*|$: Absolute Error
- $\dfrac{|x - x^*|}{|x|}$ $(x \neq 0)$: Relative Error

Which one is better?

Errors

Definition

x: the true value

x^*: an approximation to x

- $|x - x^*|$: Absolute Error
- $\dfrac{|x - x^*|}{|x|}$ $(x \neq 0)$: Relative Error

Which one is better?

Why?

Sources of Errors

- Errors in mathematical modelling: Simplifying and Assumptions

Sources of Errors

- Errors in mathematical modelling: Simplifying and Assumptions
- Blunders

Sources of Errors

- Errors in mathematical modelling: Simplifying and Assumptions
- Blunders (Prgramming Errors): Large programmes, Subprogrammes

Sources of Errors

- Errors in mathematical modelling: Simplifying and Assumptions
- Blunders (Prgramming Errors): Large programmes, Subprogrammes
- Errors in input: Errors in data transfer, uncertainties associated with measurements

Sources of Errors

- Errors in mathematical modelling: Simplifying and Assumptions
- Blunders (Prgramming Errors): Large programmes, Subprogrammes
- Errors in input: Errors in data transfer, uncertainties associated with measurements
- Machine errors by computer (Floating point arithmetic): Rounding and Chopping, Underflow and Overflow

Arithmetic

In 1985 IEEE(Institute for Electrical and Electronic Engineers) report: Binary Floating Point Arithmetic Standard 754.

Arithmetic

In 1985 IEEE(Institute for Electrical and Electronic Engineers) report: Binary Floating Point Arithmetic Standard 754.

Single, Double and Extended Precisions

Algorithms

Examining approximation procedures involving finite sequence of calculations.

- Stable: Small changes in the initial data

Algorithms

Examining approximation procedures involving finite sequence of calculations.

- Stable: Small changes in the initial data
- Unstable: Otherwise

Algorithms

Examining approximation procedures involving finite sequence of calculations.

- Stable: Small changes in the initial data
- Unstable: Otherwise
- Conditionally Stable: Stable only for certain of initial data

Convergence

$\{\alpha_n\}_{n=1}^{\infty}$, $\{\beta_n\}_{n=1}^{\infty}$: sequences

$$\lim_{n\to\infty} \beta_n = 0, \qquad \lim_{n\to\infty} \alpha_n = \alpha.$$

If $\exists\ K > 0$ s. t. $|\alpha_n - \alpha| \leq |\beta_n|$ for large n then $\{\alpha_n\}_{n=1}^{\infty}$ converges to α with the rate of convergence $O(\beta_n)$.

$$\alpha_n = \alpha + O(\beta_n)$$

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL:+82-51-510-1767 / FAX:+82-51-581-1458

PUSAN 609-735
Republic of Korea

LU factorization

$A = \mathbb{R}^{m\times n}$

$\parallel$

LU where $L \in \mathbb{R}^{m\times m}$: lower triangular matrix with diagonal entries : 1

$U \in \mathbb{R}^{m\times n}$: Echelon form of A

If $m=n$

$L \in \mathbb{R}^{n\times n}$

$U \in \mathbb{R}^{n\times n}$: upper triangular matrix

< Theory >

① Existence

② Uniqueness

③ Why?

④ How can we find?

③ Why?

$Ax = b$

$LUx = b$. Let $Ux = y$ then $LUx = \underline{Ly} = b$

y : forward substitution

대입

x : backward substitution

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL:+82-51-510-1767 / FAX:+82-51-581-1458

PUSAN 609-735
Republic of Korea

$\Updownarrow$ all eigenvalues $> 0 \Leftrightarrow$ all leading principal submatrices : positive determinant

A : symmetric positive definite $(A = A^T,\ x^T A x > 0\ \forall x \in \mathbb{R}^n)$

$\Rightarrow$ A : nonsingular $(\det(A) > 0)$

$a_{ii} > 0$

$(a_{ij})^2 < a_{ii} a_{jj}$

$A \in \mathbb{R}^{m \times n}$ $\quad \alpha \subset \{1, \cdots, m\}$
$\beta \subset \{1, \cdots, n\}$
$A(\alpha, \beta)$: submatrix
If $m = n$, $\alpha = \beta$
principal submatrix
leading principal submatrix

A : $\det(A(1:k)) \neq 0$, $k = 1, \cdots, n-1$, A : nonsingular

$\therefore$ Exists, is unique.

$A = U^T U$ (Cholesky Factorization)

proof 1.

$A = LU$ and $A^T = (LU)^T = U^T L^T$

$\therefore A = A^T \Rightarrow LU = U^T L^T \quad \therefore L = U^T \quad A = U^T U$

not true.

proof 2. By induction.

(i) $n = 1$. $\quad a = \sqrt{a}\sqrt{a} = (-\sqrt{a})(-\sqrt{a})$

$\therefore \exists!\ U \in \mathbb{R}^{n \times n}$ with positive diagonal entries s.t. $A = U^T U$.

(ii) Assume that it is true for $n-1$.

A : symmetric positive definite

$\Rightarrow$ leading principal submatrix A_{n-1} : symmetric positive definite

$$\left[\begin{array}{c|c} A_{n-1} & \\ \hline & \end{array}\right]$$

unique Cholesky factorization : $A_{n-1} = U_{n-1}^T U_{n-1}$.

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL:+82-51-510-1767 / FAX:+82-51-581-1458

PUSAN 609-735
Republic of Korea

Show : $A = U^T U$

Let $A = \begin{bmatrix} A_{n-1} & c \\ c^T & \alpha \end{bmatrix}$ where $c \in \mathbb{R}^{n-1}$, $\alpha \in \mathbb{R}$

$$\parallel \underbrace{\begin{bmatrix} U_{n-1}^T & 0 \\ r^T & \beta \end{bmatrix} \begin{bmatrix} U_{n-1} & r \\ 0 & \beta \end{bmatrix}} = U^T U$$

$$\parallel \begin{bmatrix} U_{n-1}^T U_{n-1} & U_{n-1}^T r \\ r^T U_{n-1} & r^T r + \beta^2 \end{bmatrix}$$

If $U_{n-1}^T r = c$ and $r^T r + \beta^2 = \alpha$

(i) r : unique ($\because$ U_{n-1}^T : nonsingular)

then $\beta^2 = \alpha - r^T r > 0$?

(ii) $0 < \det(A) = \det(U^T)(U) = \det(U_{n-1})^2 \beta^2$

$\therefore \beta^2 > 0$

Thus, $\exists! \ \beta > 0$

Algorithm

for $j = 1 : n$

  for $i = 1 : j-1$

    $u_{ij} = \left(a_{ij} - \sum_{k=1}^{i-1} u_{ki} u_{kj} \right) / r_{ii}$

  end

  $u_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} u_{kj}^2 \right)^{1/2}$

end

$n^3/3$ flops.

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL:+82-51-510-1767 / FAX:+82-51-581-1458

PUSAN 609-735
Republic of Korea

< QR factorization >

$A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ and $\{a_1, a_2, \cdots, a_n\}$: linearly independent

$A = [a_1\ a_2 \cdots a_n]$

$\Rightarrow$ $A = QR$ where $Q \in \mathbb{R}^{m\times n}$ with $\{q_1, q_2, \cdots q_n\}$: orthonormal

$Q = [q_1\ q_2 \cdots q_n]$

(i) $\|q_i\| = 1$ for $i = 1, \cdots, n$

(ii) $\{q_1, q_2, \cdots, q_n\}$: orthogonal set.

$\therefore$ $Q^TQ = I \in \mathbb{R}^{n\times n}$

$R \in \mathbb{R}^{n\times n}$: nonsingular upper triangular

Proof. By applying the Gram-Schmidt Process to A with normalizations

$\{a_1, a_2, \cdots, a_n\}$: linearly independent set

$\Rightarrow$ $\{q_1, q_2, \cdots, q_n\}$: orthonormal set

For each $k = 1, 2, \cdots, n$,

$$\mathrm{Span}(a_1, \cdots, a_i) = \mathrm{Span}(q_1, \cdots, q_i)$$

⌜ $\mathrm{Span}(a_1) = \mathrm{Span}(q_1)$
$\mathrm{Span}(a_1, a_2) = \mathrm{Span}(q_1, q_2)$
$\vdots$
$\mathrm{Span}(a_1, a_2, \cdots, a_n) = \mathrm{Span}(q_1, q_2, \cdots, q_n)$ ⌟

Therefore, $\exists\ r_{1i}, r_{2i}, \cdots, r_{ii}$ s.t.

$$\begin{cases} a_1 = r_{11} q_1 \\ a_2 = r_{12} q_1 + r_{22} q_2 \\ \vdots \\ a_n = r_{1n} q_1 + r_{12} q_2 + \cdots + r_{nn} q_n \end{cases}$$

$$A = [a_1\ a_2 \cdots a_n] = [q_1\ q_2 \cdots q_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = QR$$

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL:+82-51-510-1767 / FAX:+82-51-581-1458

PUSAN 609-735
Republic of Korea

Show: (i) Q has orthonormal columns.

i.e. $\{q_1, q_2 \cdots q_n\}$: orthonormal set: o.k.

(ii) R : nonsingular

$r_{ii} \neq 0$ for $i = 1, 2, \cdots, n$.

(iii) R : upper triangular : o.K

(ii) R : nonsingular

If $r_{ii} = 0$ then $a_i = r_{1i} q_1 + \cdots + r_{i-1,i} q_{i-1} + \underbrace{r_{ii} q_i}_{=0}$

$\therefore$ a_i is a linear combination of $q_1, \cdots, q_{i-1}$.

Hence $a_i \in \mathrm{Span}(q_1, \cdots, q_{i-1}) = \mathrm{Span}(a_1, \cdots, a_{i-1})$.

But $a_1, \cdots, a_{i-1}, a_i$: linearly independent, contradiction

Thus $r_{ii} \neq 0$ for $i = 1, 2, \cdots, n$.

Since R : upper triangular, R : nonsingular.

Note: Gram-Schmidt Process

$\{x_1, \cdots, x_k\}$: a basis of a subspace W

$W \subset \mathbb{R}^n$

$$v_1 = x_1$$

$$v_2 = x_2 - \left(\frac{v_1 \cdot x_2}{v_1 \cdot v_1}\right) v_1$$

$$\vdots$$

$$v_k = x_k - \left(\frac{v_1 \cdot x_k}{v_1 \cdot v_1}\right) v_1 - \left(\frac{v_2 \cdot x_k}{v_2 \cdot v_2}\right) v_2 - \cdots - \left(\frac{v_{k-1} \cdot x_k}{v_{k-1} \cdot v_{k-1}}\right) v_{k-1}$$

$W_i = \mathrm{Span}(x_1, \cdots, x_i)$ for $i = 1, \cdots, k$

Then for each $i = 1, \cdots, k$, $\{v_1, \cdots, v_i\}$: orthogonal basis for W_i.

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL : +82-51-510-1767 / FAX : +82-51-51-581-1458

PUSAN 609-735
Republic of Korea

7.1.2 Decoupling 7.1.3 The Basic Unitary Decompositions.

Theorem 7.1.3 (Schur Decomposition)

$A \in \mathbb{C}^{n\times n}$

$\exists\ Q \in \mathbb{C}^{n\times n}$, unitary s.t. $Q^H A Q = T = D + N$

where $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ and $N \in \mathbb{C}^{n\times n}$, strictly upper triangular.

proof.
(i) $n=1$: obvious
(ii) $n-1$: holds $[\ \tilde{U}^H C \tilde{U} = \tilde{T}\ ]$

Show : $Q^H A Q = T$

Find $Q = U \begin{bmatrix} 1 & 0 \\ 0 & \tilde{U} \end{bmatrix}$ U: unitary.

$$Q^H A Q = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{U}^H \end{bmatrix} \underbrace{U^H A U}_{\begin{array}{c} \| \\ \begin{array}{cc} & \\ \left[\begin{array}{cc} \lambda & w^H \\ 0 & C \end{array}\right] & \begin{array}{l} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & n-1 \end{array} & \end{array} \end{array}} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{U} \end{bmatrix} = T$$

$$\| \begin{bmatrix} \lambda & w^H \tilde{U} \\ 0 & \tilde{U}^H C \tilde{U} \end{bmatrix}$$

Question : (i) existence
(ii) $\lambda,\ w = ?$

$Ax = \lambda x \quad x \neq 0$
with $B = \lambda$.

Lemma 7.1.2

$A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{p\times p}$, $X \in \mathbb{C}^{n\times p}$

$AX = XB$, $\mathrm{rank}(X) = p$.

$$\Rightarrow \exists\ Q \in \mathbb{C}^{n\times n} \text{ unitary s.t. } Q^H A Q = T = \begin{array}{c} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{array}{l} p \\ n-p \end{array} \\ \begin{array}{cc} p & n-p \end{array} \end{array}$$

where $\lambda(T_{11}) = \lambda(A) \cap \lambda(B)$

proof. Let $X = Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ $Q \in \mathbb{C}^{n\times n}$, $R_1 \in \mathbb{C}^{p\times p}$

: QR factorization of X.

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL : +82-51-510-1767 / FAX : +82-51-51-581-1458

PUSAN 609-735
Republic of Korea

Let $Q^H A Q = \begin{array}{cc} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} & \begin{array}{c} p \\ n-p \end{array} \end{array} = T$

$$\begin{array}{c} \begin{array}{cc} p & n-p \end{array} \\ \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \end{array} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \overset{T = Q^H A Q}{=} Q^H A Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \overset{X = QR}{=} Q^H A X \overset{AX = XB}{=} Q^H X B \overset{X = QR}{=} Q^H Q \begin{bmatrix} R_1 \\ 0 \end{bmatrix} B$$

$$= \begin{bmatrix} R_1 \\ 0 \end{bmatrix} B.$$

Since R_1 : nonsingular, $T_{21} R_1 = 0$ and $T_{11} R_1 = R_1 B$.

$R_1^{-1} T_{11} R_1 = B$

T_{11} is similar to B.

$\therefore$ $T_{21} = 0$ and $\lambda(T_{11}) = \lambda(B)$

Show : $\lambda(T_{11}) = \lambda(A) \cap \lambda(B)$

Suppose $\lambda(A) = \lambda(T) = \lambda(T_{11}) \cup \lambda(T_{22})$: we need to prove

O.K. $\Rightarrow$ $\lambda(A) \cap \lambda(B) = \left(\lambda(T_{11}) \cup \lambda(T_{22})\right) \cap \lambda(T_{11})$

$= \lambda(T_{11})$.

Lemma 7.1.1.

$$T \in \mathbb{C}^{n \times n} \quad \Rightarrow \quad \lambda(T) = \lambda(T_{11}) \cup \lambda(T_{22}).$$

$$\| \quad \begin{array}{cc} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} & \begin{array}{c} p \\ q \end{array} \\ \begin{array}{cc} p & q \end{array} & \end{array}$$

proof.

$$Tx = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{where } x_1 \in \mathbb{C}^p \quad x_2 \in \mathbb{C}^q$$

(i) $x_2 \neq 0$

$T_{22} x_2 = \lambda x_2 \Rightarrow \lambda \in \lambda(T_{22})$

(ii) $x_2 = 0$

$T_{11} x_1 = \lambda x_1 \Rightarrow \lambda \in \lambda(T_{11})$

$\therefore$ $\lambda(T) \subset \lambda(T_{11}) \cup \lambda(T_{22})$

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL : +82-51-510-1767 / FAX : +82-51-51-581-1458

PUSAN 609-735
Republic of Korea

General version of Lemma 7.1.5

Theorem 7.1.6 (Block Diagonal Decomposition)

$A \in \mathbb{C}^{n\times n}$

(i) $Q^H A Q = T = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1q} \\ & T_{22} & \cdots & T_{2q} \\ 0 & & \ddots & \vdots \\ & & & T_{qq} \end{bmatrix}$. T_{ii} : square for $i=1,\cdots,q$

(ii) $\lambda(T_{ii})$ and $\lambda(T_{jj})$: disjonit whenever $i \neq j$

$\Rightarrow \exists X$: nonsingular s.t. $X^{-1} A X = \operatorname{diag}(T_{11},\cdots,T_{qq})$

Sketch of proof.

By induction

(i) 2x2 block : $Q^H A Q = T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$ $\lambda(T_{11}) \cap \lambda(T_{22}) = \emptyset$. T_{11}, T_{22} : square

then $\exists Y$ s.t. $Y^{-1} T Y = \operatorname{diag}(T_{11}, T_{22})$

(ii) 3x3 block : $Q^H A Q = T = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ 0 & T_{22} & T_{23} \\ 0 & 0 & T_{33} \end{bmatrix}$

then $\exists Y_1$ s.t. $Y_1^{-1} T Y_1 = \begin{bmatrix} T_{11} & T_{12} & 0 \\ 0 & T_{22} & 0 \\ 0 & 0 & T_{33} \end{bmatrix}$

$\exists Y_2$ s.t. $Y_2^{-1}(Y_1^{-1} T Y_1) Y_2 = \begin{bmatrix} T_{11} & & 0 \\ & T_{22} & \\ 0 & & T_{33} \end{bmatrix}$

(iii) 4x4 block

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL: +82-51-510-1767 / FAX: +82-51-581-1458

PUSAN 609-735
Republic of Korea

Lecture 4. The Singular Value Decomposition

A Geometric Observation

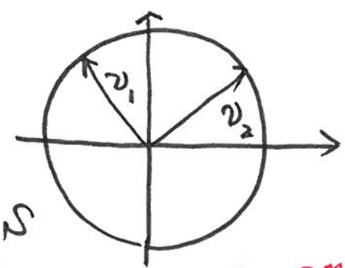
The image of the unit sphere under any $m \times n$ matrix is a hyperellipse.

m-dimensional generalization of an ellipse.

Consider the unit sphere in $\mathbb{R}^n$.

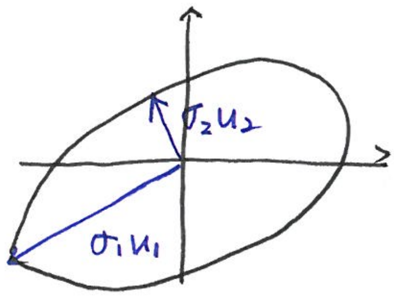
orthogonal directions $u_1, \cdots, u_m \in \mathbb{R}^m$ $\|u_i\| = 1$.

$\sigma_1, \cdots, \sigma_m$ (possibly zero)



S

AS



Reduced SVD $A \in \mathbb{C}^{m\times n}$ $(m \geq n)$

$$A v_j = \sigma_j u_j \qquad 1 \leq j \leq n. \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$$

$$A \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} =$$

$$= \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

$$AV = \hat{U} \hat{\Sigma}$$

$$A = \underset{m\times n}{\hat{U}} \; \underset{n\times n}{\hat{\Sigma}} \; \underset{n\times n}{V^*} \quad (V: \text{unitary}).$$

(A: assumed full rank n)

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL: +82-51-510-1767 / FAX: +82-51-581-1458

PUSAN 609-735
Republic of Korea

Full SVD

① $A \in \mathbb{C}^{m \times n}$

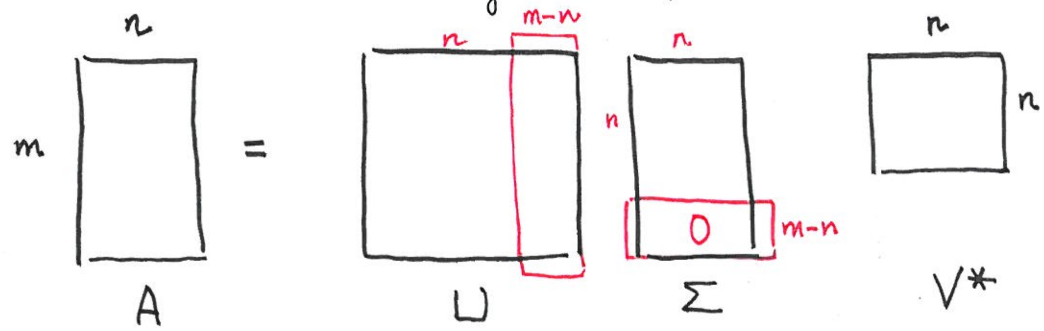
② $m \geq n$

$A = U \Sigma V^*$

U: $m \times m$ unitary

Σ: $m \times n$ diagonal with positive real entries

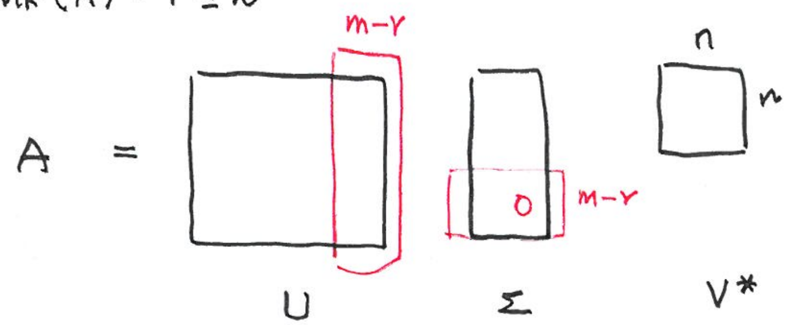
V: $n \times n$ unitary



n : singular vectors of A

③ A is not assumed to have full rank n.
(A is rank-deficient. $\operatorname{rank}(A) \leq n$)

Let $\operatorname{rank}(A) = r \leq n$



Note: $m-n \leq m-r$ ∴ more "0" rows.

DEPARTMENT OF MATHEMATICS
PUSAN NATIONAL UNIVERSITY
TEL : +82-51-510-1767 / FAX : +82-51-581-1458

PUSAN 609-735
Republic of Korea

Theorem

① The nonzero singular values of A

= the square roots of the nonzero eigenvalues of A^*A or AA^*.

$A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^* U^* U\Sigma V^* = V(\underbrace{\Sigma^*\Sigma}_{\sigma_1^2 \cdots \sigma_r^2})V^*$

② If $A = A^*$ then the singular values of A are the absolute values of the eigenvalues of A.

③ $A \in \mathbb{C}^{m\times m}$, $\quad |\det(A)| = \prod_{i=1}^{m} \sigma_i$

$$|\det(A)| = |\det(U\Sigma V^*)| = |\det(U)|\,|\det(\Sigma)|\,|\det(V^*)| = |\det(\Sigma)| = \prod_{i=1}^{m} \sigma_i$$