Lesson 14

Chapter 7

Symmetric Matrices and Quadratic Forms

▸ Diagonalization of Symmetric Matrices ▸ Quadratic Forms ▸ Constrained Optimization

▸ The Singular Value Decomposition ▸ Applications to Image Processing and Statistics

Spectral decomposition:

Spectrum of A

Let: A a symmetric $n \times n$ matrix with eigenvalues: $\lambda_1, \cdots, \lambda_n$ and corresponding orthononal eigenvectors: $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$.

Then:

$$A = PDP^{-1} = PDP^T = [\boldsymbol{u}_1 \ \dots \ \boldsymbol{u}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} = [\lambda_1 \boldsymbol{u}_1 \ \dots \ \lambda_n \boldsymbol{u}_n] \begin{bmatrix} \boldsymbol{u}_1^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix}$$

$$= \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T + \cdots + \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T$$

Spectral decomposition of A

Også kaldet: egenværdi-dekomposition

$n \times n$ matrix with rank 1

Projection matrix on subspace spanned by $\boldsymbol{u}_i$:

$(\boldsymbol{u}_i \boldsymbol{u}_i^T)\boldsymbol{x} = proj_{\boldsymbol{u}_i} \boldsymbol{x}$

Geometric interpretation of Principal Axes

$\boldsymbol{x}^T A\boldsymbol{x} = c;\ \ \boldsymbol{x} \in \mathbb{R}^2,\ \ c \in \mathbb{R},\ \ A$ symmetric 2×2 matrix

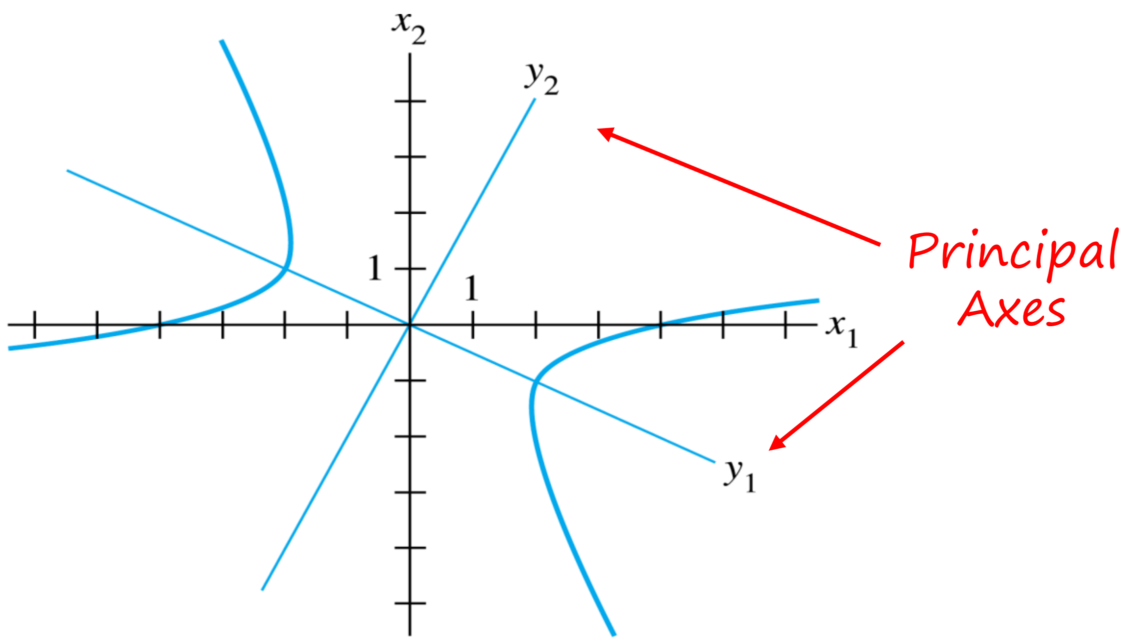
$$A = \begin{bmatrix} a & d \\ d & -b \end{bmatrix} = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix};\ \ c = 16$$

$$\boldsymbol{x}^T A\boldsymbol{x} = c \ \ \Leftrightarrow \ \ x_1^2 - 8x_1x_2 - 5x_2^2 = 16$$

$$\lambda_1 = 3:\ \boldsymbol{u}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix} \qquad \lambda_2 = -7:\ \boldsymbol{u}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$P = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

$$\boldsymbol{y} = P^{-1}\boldsymbol{x} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2x_1 - x_2 \\ x_1 + 2x_2 \end{bmatrix}$$



(b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

$3y_1^2 - 7y_2^2 = 16$

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{y}^T D\boldsymbol{y} = c \ \ \Leftrightarrow \ \ 3y_1^2 - 7y_2^2 = 16 \ \Leftrightarrow \ \frac{y_1^2}{\left(4/\sqrt{3}\right)^2} - \frac{y_2^2}{\left(4/\sqrt{7}\right)^2} = 1$$

7.4 The Singular Value Decomposition

Matrix diagonalization: $\boldsymbol{A = PDP^{-1}}$

$P = [\boldsymbol{u}_1\ \boldsymbol{u}_2\ \cdots\ \boldsymbol{u}_n]$ and $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$, where λ_i and $\boldsymbol{u}_i$ are the eigenvalues and -vectors of A

- **The Diagonalization Theorem (5.5):**
  $n \times n$ matrix A diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors

- **The Spectral Theorem (7.3):**
  $n \times n$ symmetric matrix A is orthogonally diagonalizable
  → Spectral decomposition: $A = PDP^{-1} = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T$

- **Not all** matrices can be factored as: $A = PDP^{-1}$

- **Any** $m \times n$ matrix can be factored as: $A = QDP^{-1}$
  → Singular value decomposition → Very helpfull in many computer calculations

Optimatization problem:

A linear transformation: $\boldsymbol{x} \mapsto A\boldsymbol{x}$ where A is an $m \times n$ matrix.

Which unit vector $\boldsymbol{x}$ maximize $A\boldsymbol{x}$?

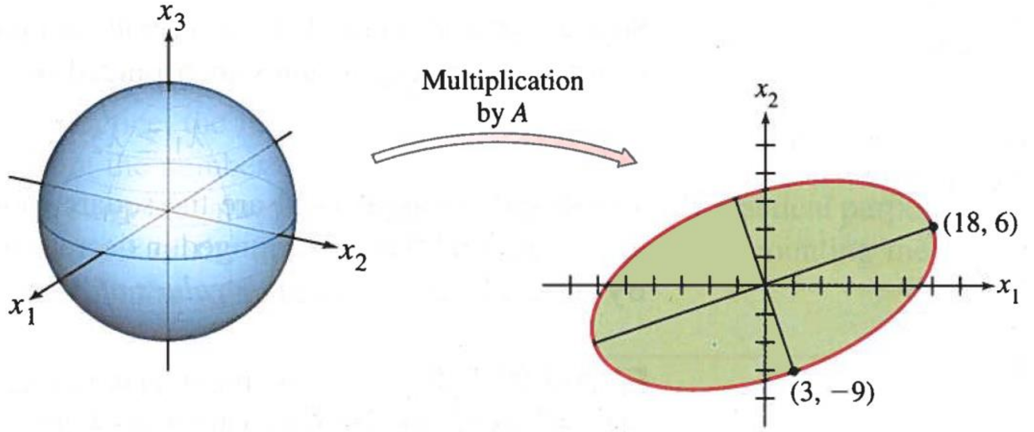
$$||A\boldsymbol{x}||^2 = (A\boldsymbol{x})^T A\boldsymbol{x} = \boldsymbol{x}^T A^T A\boldsymbol{x} = \boldsymbol{x}^T (A^T A)\boldsymbol{x}$$

→ Quadratic form ($A^T A$ symmetric) with constrain $||\boldsymbol{x}||$=1

→ Theorem 7.6: Max. value = λ_{max} in direction of the corresponding eigenvector

Ex 1 Linear transformation: $\boldsymbol{x} \mapsto A\boldsymbol{x}$ with $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

Unit sphere $\|\boldsymbol{x}\| = 1$ in $\mathbb{R}^3 \mapsto$ Ellipse in $\mathbb{R}^2$



$$A^T A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 360; \ \ \lambda_2 = 90; \ \ \lambda_3 = 0$

Unit eigenvectors: $\boldsymbol{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}; \ \ \boldsymbol{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}; \ \ \boldsymbol{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$ → $\|A\boldsymbol{x}\|_{max} = \lambda_1 = 360$, when $\boldsymbol{x} = \boldsymbol{v}_1$

$A\boldsymbol{v}_1 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \end{bmatrix}$ → the point farthest from the origin

$\|A\boldsymbol{v}_1\| = \sqrt{\lambda_1} = \sqrt{360} \approx 19{,}0$

Singular Values:

$m \times n$ matrix $A \Rightarrow A^T A$ symmetric $\Leftrightarrow A^T A$ orthogonally diagonizable:

$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ an orthonornal basis for $\mathbb{R}^n$ of eigenvectors of $A^T A$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ the corresponding (ordered set of) eigenvalues

$$\rightarrow \ \|A\boldsymbol{v}_i\|^2 = (A\boldsymbol{v}_i)^T A\boldsymbol{v}_i = \boldsymbol{v}_i{}^T A^T A\boldsymbol{v}_i = \boldsymbol{v}_i{}^T(\lambda_i \boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i{}^T \boldsymbol{v}_i = \lambda_i \geq 0$$

$$\boxed{\rightarrow \ \|A\boldsymbol{v}_i\| = \sigma_i = \sqrt{\lambda_i} \longleftarrow \text{Singular Values}}$$

Computation exercise:

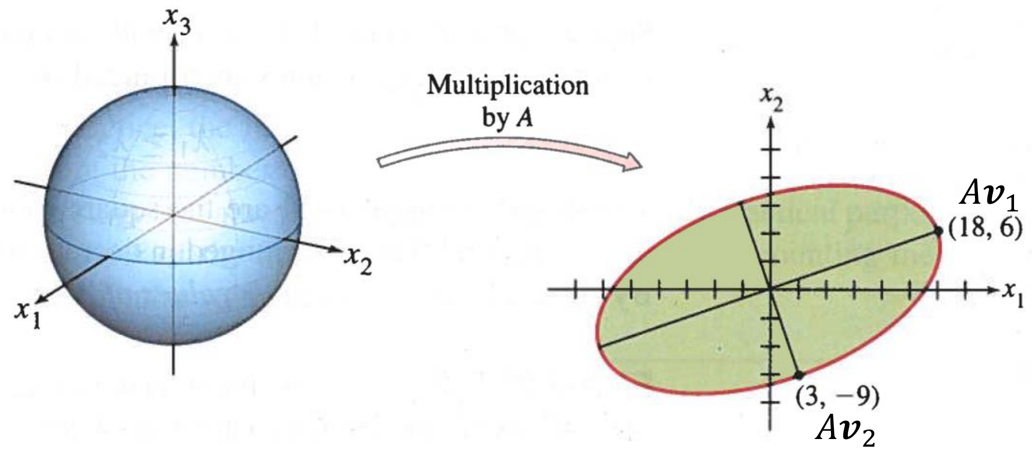
Here's two matrices

$$A_1 = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$$

Compute the singular values of A_1 and A_2.

Ex 2 Linear transformation: $\boldsymbol{x} \mapsto A\boldsymbol{x}$ with $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

Unit sphere $\|\boldsymbol{x}\| = 1$ in $\mathbb{R}^3 \mapsto$ Ellipse in $\mathbb{R}^2$



$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 360$; $\lambda_2 = 90$; $\lambda_3 = 0$ → Singular Values: $\sigma_1 = \sqrt{360} \approx 19{,}0$; $\sigma_2 = \sqrt{90} \approx 9{,}5$; $\sigma_3 = 0$

Unit eigenvectors: $\boldsymbol{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$; $\boldsymbol{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$; $\boldsymbol{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$

$\boxed{\text{OBS: } A\boldsymbol{v}_1 \perp A\boldsymbol{v}_2 \text{ and } \boldsymbol{v}_3 \in Nul\ A}$

$$A\boldsymbol{v}_1 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \end{bmatrix}$$

$\|A\boldsymbol{v}_1\| = \sigma_1 = \sqrt{360} \approx 19{,}0$

$$A\boldsymbol{v}_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}$$

$\|A\boldsymbol{v}_2\| = \sigma_2 = \sqrt{90} \approx 9{,}5$

$$A\boldsymbol{v}_3 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\|A\boldsymbol{v}_3\| = \sigma_3 = 0$

Theorem 7.9

Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A^TA arranged so that the corresponding eigenvalues of A^TA satisfy $\lambda_1 \geq \cdots \geq \lambda_n$, and suppose A has r nonzero singular values.

Then $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is an orthogonal basis for col A and rank $A = r$.

Numerical note:

Most reliable way to estimate the rank of a large matrix A is to count the number of nonzero singular values. Extremely small nonzero singular values are assumed to be zero.

Diskussion:

Givet A er en 5*7 matrix:

Hvor mange singulærværdier (større end 0) kan A maksimalt have?

<u>Theorem 7.10</u>: Singular Value Decomposition (SVD)

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix Σ

$$\Sigma = \begin{array}{c@{}l} \begin{array}{cc} \overbrace{}^{r} & \overbrace{}^{n-r \text{ columns}} \end{array} & \\ \left[\begin{array}{cc} D & 0 \\ 0 & 0 \end{array}\right] & \begin{array}{l} \} \ r \text{ rows} \\ \} \ m-r \text{ rows} \end{array} \end{array}$$

for which the diagonal entries in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and there exists an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that

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

Left singular vectors *Right singular vectors*

Ex 3 Singular value decomposition of $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

Step 1: Orthogonal diagonalization of $A^TA = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$ → $\begin{cases} \text{Eigenvalues: } \lambda_1 = 360; \quad \lambda_2 = 90; \quad \lambda_3 = 0 \\ \text{Eigenvectors: } \boldsymbol{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}; \quad \boldsymbol{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}; \quad \boldsymbol{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix} \end{cases}$

Step 2: Set up $\boldsymbol{V}$ and $\boldsymbol{\Sigma}$ → Singular Values (in decreasing order): $\sigma_1 = \sqrt{360} = 6\sqrt{10}; \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}; \quad \sigma_3 = 0$

$$\to V = [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}; \quad D = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}; \quad \Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

Right singular vectors

Step 3: Construct $\boldsymbol{U}$ → Rank $A = 2$ → $\boldsymbol{u}_1 = \frac{A\boldsymbol{v}_1}{\|A\boldsymbol{v}_1\|} = \frac{A\boldsymbol{v}_1}{\sigma_1} = \frac{1}{6\sqrt{10}}\begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}; \quad \boldsymbol{u}_2 = \frac{A\boldsymbol{v}_2}{\|A\boldsymbol{v}_2\|} = \frac{A\boldsymbol{v}_2}{\sigma_2} = \frac{1}{3\sqrt{10}}\begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$

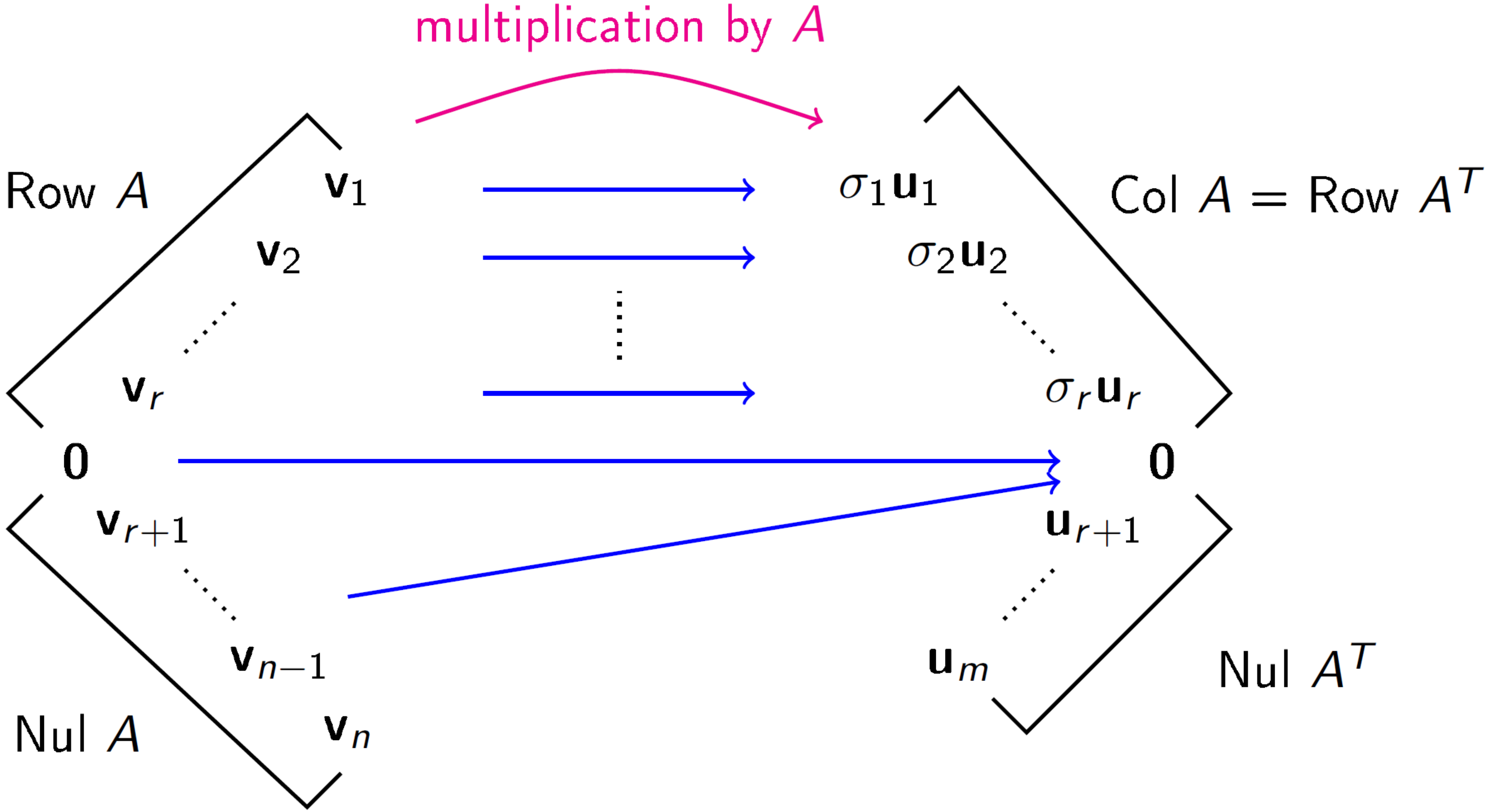
$$\to U = [\boldsymbol{u}_1\ \boldsymbol{u}_2] = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}$$ ($\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ basis for $\mathbb{R}^2$ → no additional vectors needed for U)

Left singular vectors

→ **Singular Value Decomposition**: $A = U\Sigma V^T = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$

The four fundamental subspaces of the $m \times n$ matrix $A = U\Sigma V^T$

$Rank\ A = r$ $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ = Right singular vectors $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m\}$ = Left singular vectors



The Invertible Matrix Theorem

Let A be a square $n\ x\ n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.

b. A is row equivalent to the $n \times n$ identity matrix.

c. A has n pivot positions.

d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

g. The equation $A\mathbf{x} = \mathbf{b}$ has ~~at least one~~ a unique solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$.

i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C such that $CA = I$.

k. There is an $n \times n$ matrix D such that $AD = I$.

l. A^T is an invertible matrix.

The Invertible Matrix Theorem - continued

Let A be a square $n\ x\ n$ matrix. Then the following statements are equivalent.
That is, for a given A, the statements are either all true or all false.

m. The columns of A form a basis of $\mathbb{R}^n$.

n. Col $A = \mathbb{R}^n$.

o. dim Col $A = n$.

p. Rank $A = n$.

q. Nul $A = \{\mathbf{0}\}$.

r. dim Nul A=0.

s. The number 0 is not a eigenvalue of A.

t. The determinant of A is not 0.

New

u. $(\text{Col } A)^{\perp} = \{\mathbf{0}\}$.

v. $(\text{Nul } A)^{\perp} = \mathbb{R}^n$.

w. Row $A = \mathbb{R}^n$.

x. A has n nonzero singular values.

OBS!
Due to (l): A invertible $\Leftrightarrow A^T$ invertible
and Row A = Col A^T:

- All statement could also be stated for A^T
- All statements on Col A could also be stated on Row A

Diskussion:

Givet A er en 3*3 matrix med singulærværdier $\sigma_1 = 4$, $\sigma_2 = 3$, $\sigma_3 = 0$.

- Er A mon invertibel? Hvorfor/hvorfor ikke?
- Lad nu σ_3 i stedet være 0.001. Diskuter A's invertibilitet

Ex 4 An $m \times n$ matrix $A = U\Sigma V^T$ with Rank $A = r < \max(m, n)$

→ Σ contains rows and/or –columns of zeros

$U = [U_r \;\; U_{m-r}]$, where $U_r = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_r]$ $\qquad$ $V = [V_r \;\; V_{n-r}]$, where $V_r = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_r]$

$$\rightarrow\; A = U\Sigma V^T = [U_r \;\; U_{m-r}]\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} {V_r}^T \\ {V_{n-r}}^T \end{bmatrix} = [U_r \;\; U_{m-r}]\begin{bmatrix} D{V_r}^T \\ 0 \end{bmatrix} = U_r D {V_r}^T$$

D invertible ← nonzero diagonal elements

Reduced singular value decomposition

$$\rightarrow\; A^+ = V_r D^{-1} {U_r}^T$$

The pseudoinverse of A (OBS: $AA^+ \neq I$)

Inconsistent equations: $A\boldsymbol{x} = \boldsymbol{b}$

Define: $\hat{\boldsymbol{x}} = A^+\boldsymbol{b} = V_r D^{-1} {U_r}^T \boldsymbol{b}$

→ $A\hat{\boldsymbol{x}} = U_r D {V_r}^T V_r D^{-1} {U_r}^T \boldsymbol{b} = U_r D D^{-1} {U_r}^T \boldsymbol{b} = U_r {U_r}^T \boldsymbol{b} = \hat{\boldsymbol{b}} = proj_{Col\,A}\boldsymbol{b}$

→ $\hat{\boldsymbol{x}} = A^+\boldsymbol{b}$ is a least-square solution to $A\boldsymbol{x} = \boldsymbol{b}$ (the solution with smallest length)

OPGAVE 1

For matricen A:

$$A = \begin{bmatrix} 0.75 & -0.25 & 0.25 \\ 0.50 & -0.50 & -0.50 \\ -0.50 & 1.00 & 0.45 \\ 0.25 & 0.25 & 0.75 \end{bmatrix}$$

1. Bestem singulærværdi-dekompositionen A = USV^T
2. Betragt singulærværdierne σ_i og afgør hvilken rank A har
3. Hvad kan vi sige om lineært uafhængige søjler i A?

4. Erstat A(3,3) med værdien 1.45 og gentag step 1-3
5. Sæt nu σ_3 i S matricen til 0 og beregn $\hat{A} = US_{trunc}V^T$
6. Hvilken effekt får det? Hvordan adskiller $\hat{A}$ sig fra A?

OPGAVE 2

For en matrix A er der lavet SVD resulterende i:

$$\Sigma = \begin{bmatrix} 3.4 & 0 & 0 \\ 0 & 1.6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Hvad kan vi sige om lineært uafhængige søjler i A?
- Hvad kan vi udlede omkring løsning af ligningssystemet: A**x** = **b**

- Beregn conditionstallet hvis $\sigma_3 = 0.1$ i stedet for 0

Anvendelser af Singular Value Decomposition (SVD):

- Støjfjernelse
- Systemidentifikation
- Estimation af hvor ”sundt” (læs: ortogonalt) er datasæt er
- Identifikation af redundant information i en datamatrix A
- Beregning af PCA …

  .

  .

  .

  .

  .