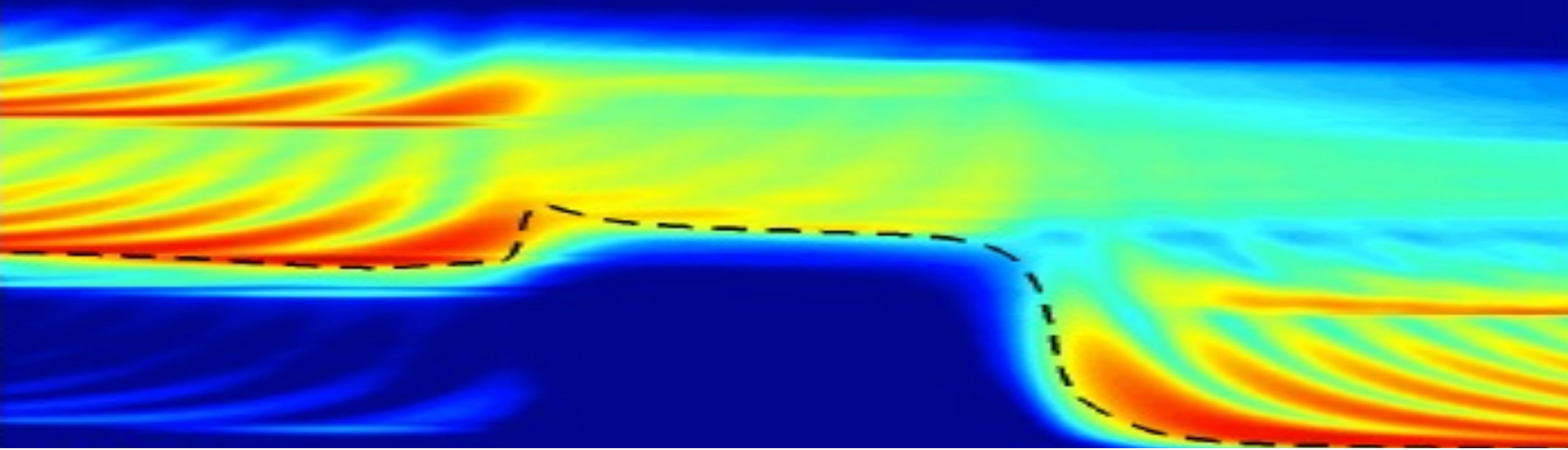


EECS 16B
Designing Information Devices and Systems II
Lecture 21

Prof. Yi Ma
Department of Electrical Engineering and Computer Sciences, UC Berkeley,
yima@eecs.berkeley.edu

Outline

- Motivations for Singular Value Decomposition
  - Least Squares and Minimum Norm Solution
  - Identifying Low-dim Linear Subspace
- Singular Value Decomposition (SVD)
  - Algorithm
  - Example
  - Theorem (with proof)

Least-Squares vs Minimum-Norm Solutions

Moore-Penrose pseudo inverse of $A \in \mathbb{R}^{m \times n}$: $\vec{y} = A\vec{x}, \quad \vec{x} = A^\dagger \vec{y}.$

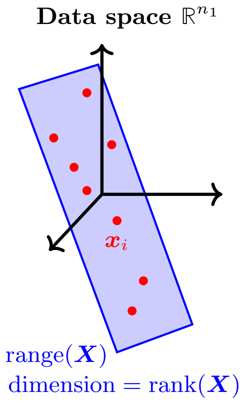
$$m \geq n \text{ and } \operatorname{rank}(A) = n : \quad A^\dagger = (AA^\top)^{-1}A^\top$$

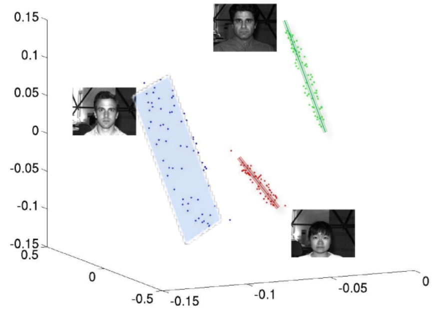
$$m \leq n \text{ and } \operatorname{rank}(A) = m : \quad A^\dagger = A^\top (AA^\top)^{-1}$$

$A \in \mathbb{R}^{m \times n}$ not full column or row rank? Moore-Penrose pseudo inverse

Low-Dim Structures in High-Dim Data

Principal Component Analysis: Finding one linear subspace

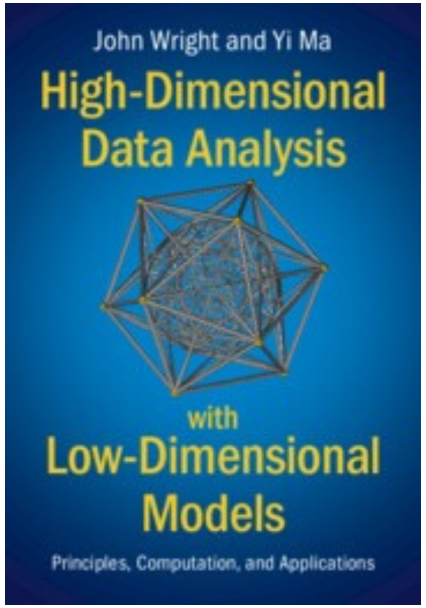




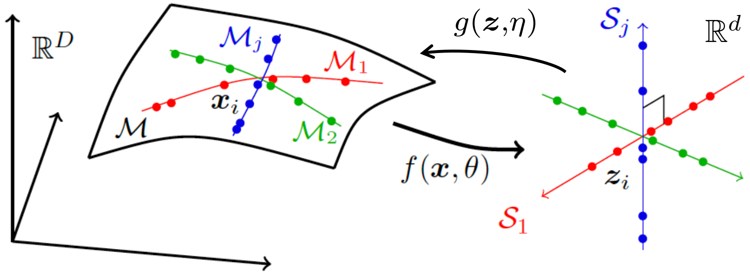
Compressive Sensing: Finding multiple low-dim linear structures

- Solving under-determined systems of linear equations
- Low-rank matrix approximation or recovery

Deep Learning: Finding non-linear low-dimensional structures



A mixture of low-dim submanifolds



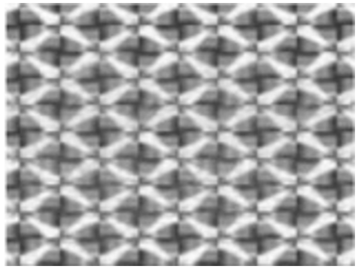
A mixture of incoherent subspaces

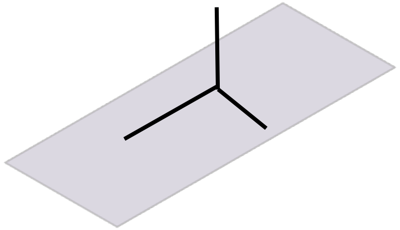
EECS 208: Computational Principles for High-Dimensional Data Analysis

(from SVD/PCA, to Generalized PCA, Robust PCA, Nonlinear PCA, and to Deep Networks…)

Identifying a Low-dim Linear Subspace

Given $X = [\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n] \in \mathbb{R}^{m \times n}$, find a low rank L : $\min_L \|X - L\|_F^2, \quad \text{s.t.} \quad \text{rank}(L) \leq r.$





$$X = [\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n] \in \mathbb{R}^{m \times n}$$

Singular Value Decomposition (SVD)

Given $A \in \mathbb{R}^{m\times n}$ with $\text{rank}(A) = r$, we like to decompose it into a special outer-product form:

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^\top + \sigma_2 \vec{u}_2 \vec{v}_2^\top + \cdots + \sigma_r \vec{u}_r \vec{v}_r^\top$$

$\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r$ orthonormal in $\mathbb{R}^m$

$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ orthonormal in $\mathbb{R}^n$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

Singular Value Decomposition (SVD)

Given $A \in \mathbb{R}^{m \times n}$ with $\mathrm{rank}(A) = r$, we like to decompose it into a special matrix form:

$$A = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \\ \vdots \\ \vec{v}_r^\top \end{bmatrix}$$

$$U = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r] \text{ orthogonal}$$

$$V = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r] \text{ orthogonal}$$

$$\Sigma = \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_r\} > 0$$

Singular Value Decomposition (SVD)

Claim: $A^\top A \in \mathbb{R}^{n \times n}$ all eigenvalues are non-negative and eigenvectors are orthogonal.

Proof:

Singular Value Decomposition (SVD)

Claim: $\operatorname{rank}(A^\top A) = \operatorname{rank}(A) = r$ hence r eigenvalues of $A^\top A$ are positive.

Proof:

Singular Value Decomposition (Algorithm)

Algorithm: given $A \in \mathbb{R}^{m\times n}$ with $\mathrm{rank}(A) = r$, start with. $A^\top A \in \mathbb{R}^{n\times n}$:

Singular Value Decomposition (Example)

Singular Value Decomposition (Theorem)

Theorem: given $A \in \mathbb{R}^{m\times n}$ with $\mathrm{rank}(A) = r$, let $A^\top A = \sum_{i=1}^{r} \lambda_i \vec{v}_i \vec{v}_i^\top$ and $\sigma_i = \sqrt{\lambda_i}$, $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i \in \mathbb{R}^m,\ i = 1, \ldots, r$. Then we have $U = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r]$ orthogonal, and

$$A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top = U\Sigma V^\top$$

$$\Sigma = \mathrm{diag}\{\sigma_1, \ldots, \sigma_r\} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_r \end{bmatrix}$$

Proof:

Singular Value Decomposition (Theorem)

Theorem: given $A \in \mathbb{R}^{m \times n}$ with $\mathrm{rank}(A) = r$, let $A^\top A = \sum_{i=1}^{r} \lambda_i \vec{v}_i \vec{v}_i^\top$ and $\sigma_i = \sqrt{\lambda_i}$, $\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i \in \mathbb{R}^m,\ i = 1, \ldots, r$. Then we have $U = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r]$ $\mathrm{orthogonal}$, and

$$A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top = U \Sigma V^\top$$

Proof: