Singular Value Decomposition

CS 205A:
Mathematical Methods for Robotics, Vision, and Graphics

Justin Solomon

Understanding the Geometry of $A \in \mathbb{R}^{m \times n}$

Critical points of the ratio:

$$R(\vec{x}) = \frac{\|A\vec{x}\|_2}{\|\vec{x}\|_2}$$

Understanding the Geometry of $A \in \mathbb{R}^{m \times n}$

Critical points of the ratio:

$$R(\vec{x}) = \frac{\|A\vec{x}\|_2}{\|\vec{x}\|_2}$$

- $R(\alpha\vec{x}) = R(\vec{x}) \implies$ take $\|\vec{x}\|_2 = 1$
- $R(\vec{x}) \geq 0 \implies$ study $R^2(\vec{x})$ instead

Once Again...

Critical points satisfy

$$A^\top A\vec{x}_i = \lambda_i\vec{x}_i$$

Once Again...

Critical points satisfy

$$A^\top A\vec{x}_i = \lambda_i\vec{x}_i$$

Properties:

- $\lambda_i \geq 0 \; \forall i$

Once Again...

Critical points satisfy

$$A^\top A\vec{x}_i = \lambda_i \vec{x}_i$$

Properties:

- $\lambda_i \geq 0 \, \forall i$
- Basis is full and orthonormal

Geometric Question

How is $\vec{x}_i$ related to the geometry of A rather than that of $A^\top A$?

Geometric Question

How is $\vec{x}_i$ related to the geometry of A rather than that of $A^\top A$?

Object of study: $\vec{y}_i = A\vec{x}_i$

Observation

Lemma

Either $\vec{y}_i = \vec{0}$ or $\vec{y}_i$ is an eigenvector of $AA^\top$ with $\|\vec{y}_i\| = \sqrt{\lambda_i}\|\vec{x}_i\|$.

Corresponding Eigenvalues

$$
\begin{aligned}
k &= \text{ number of } \lambda_i > 0 \\
A^\top A\vec{x}_i &= \lambda_i \vec{x}_i \\
AA^\top \vec{y}_i &= \lambda_i \vec{y}_i \\
\bar{U} \in \mathbb{R}^{n\times k} &= \text{ matrix of } \vec{x}_i\text{'s} \\
\bar{V} \in \mathbb{R}^{m\times k} &= \text{ matrix of } \vec{y}_i\text{'s}
\end{aligned}
$$

Corresponding Eigenvalues

$$
\begin{aligned}
k &= \text{ number of } \lambda_i > 0 \\
A^\top A\vec{x}_i &= \lambda_i \vec{x}_i \\
AA^\top \vec{y}_i &= \lambda_i \vec{y}_i \\
\bar{U} \in \mathbb{R}^{n\times k} &= \text{ matrix of } \vec{x}_i\text{'s} \\
\bar{V} \in \mathbb{R}^{m\times k} &= \text{ matrix of } \vec{y}_i\text{'s}
\end{aligned}
$$

Lemma

$$\bar{U}^\top A \bar{V} \vec{e}_i = \sqrt{\lambda_i}\vec{e}_i$$

Observation

Lemma

$\bar{U}^\top A \bar{V} \vec{e}_i = \sqrt{\lambda_i}\vec{e}_i$

$$\bar{\Sigma} \equiv \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_k})$$

Observation

Lemma

$\bar{U}^\top A \bar{V} \vec{e}_i = \sqrt{\lambda_i} \vec{e}_i$

$$\bar{\Sigma} \equiv \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_k})$$

Corollary

$\bar{U}^\top A \bar{V} = \bar{\Sigma}$

Completing the Basis

Add $\vec{x}_i$ with $A^\top A\vec{x}_i = \vec{0}$ and $\vec{y}_i$ with $AA^\top\vec{y}_i = \vec{0}$

Completing the Basis

Add $\vec{x}_i$ with $A^\top A\vec{x}_i = \vec{0}$ and $\vec{y}_i$ with $AA^\top\vec{y}_i = \vec{0}$

$$\bar{U} \in \mathbb{R}^{m\times k}, \bar{V} \in \mathbb{R}^{n\times k} \mapsto$$
$$U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$$

Completing the Basis

Add $\vec{x}_i$ with $A^\top A\vec{x}_i = \vec{0}$ and $\vec{y}_i$ with $AA^\top\vec{y}_i = \vec{0}$

$$\bar{U} \in \mathbb{R}^{m\times k}, \bar{V} \in \mathbb{R}^{n\times k} \mapsto$$
$$U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$$

$$\Sigma_{ij} \equiv \begin{cases} \sqrt{\lambda_i} & i = j \text{ and } i \leq k \\ 0 & \text{otherwise} \end{cases}$$

Singular Value Decomposition

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

SVD Vocabulary

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

- Left singular vectors: Columns of U; span $\operatorname{col} A$
- Right singular vectors Columns of V; span $\operatorname{row} A$
- Singular values: Diagonal σ_i of Σ; sort $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$.

Geometry of Linear Transformations

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

1. Rotate $(V^\top)$
2. Scale (Σ)
3. Rotate (U)

Computing SVD: Simple Strategy

1. Columns of V are eigenvectors of $A^\top A$
2. $AV = U\Sigma \implies$ columns of U corresponding to nonzero singular values are normalized columns of AV
3. Remaining columns of U satisfy $AA^\top \vec{u}_i = \vec{0}$.

Computing SVD: Simple Strategy

1. Columns of V are eigenvectors of $A^\top A$
2. $AV = U\Sigma \implies$ columns of U corresponding to nonzero singular values are normalized columns of AV
3. Remaining columns of U satisfy $AA^\top \vec{u}_i = \vec{0}$.

$\exists$ **more specialized methods!**

Solving Linear Systems with $A = U\Sigma V^\top$

$$
\begin{aligned}
A\vec{x} &= \vec{b} \\
\implies U\Sigma V^\top \vec{x} &= \vec{b} \\
\implies \vec{x} &= V\Sigma^{-1}U^\top \vec{b}
\end{aligned}
$$

Solving Linear Systems with $A = U\Sigma V^\top$

$$A\vec{x} = \vec{b}$$

$$\implies U\Sigma V^\top \vec{x} = \vec{b}$$

$$\implies \vec{x} = V\Sigma^{-1}U^\top \vec{b}$$

What is Σ^{-1}?

Uniting Short/Tall Matrices

$$\text{minimize } \|\vec{x}\|_2^2$$
$$\text{such that } A^\top A\vec{x} = A^\top \vec{b}$$

Simplification

$$A^\top A = V\Sigma^2 V^\top$$

Simplification

$$A^\top A = V\Sigma^2 V^\top$$

$$A^\top A\vec{x} = A^\top\vec{b} \Leftrightarrow \Sigma\vec{y} = \vec{d}$$

$$\vec{y} \equiv V^\top\vec{x}$$

$$\vec{d} \equiv U^\top\vec{b}$$

Resulting Optimization

$$\text{minimize } \|\vec{y}\|_2^2$$
$$\text{such that } \Sigma\vec{y} = \vec{d}$$

Solution

$$\Sigma^+_{ij} \equiv \begin{cases} 1/\sigma_i & i = j, \sigma_i \neq 0, \text{ and } i \leq k \\ 0 & \text{otherwise} \end{cases}$$

$$\implies \vec{y} = \Sigma^+ \vec{d}$$

$$\implies \vec{x} = V\Sigma^+ U^\top \vec{b}$$

Pseudoinverse

$$A^{+} = V\Sigma^{+}U^{\top}$$

Pseudoinverse Properties

- A **square** and **invertible** $\implies A^+ = A^{-1}$
- A **overdetermined** $\implies A^+\vec{b}$ gives least-squares solution to $A\vec{x} \approx \vec{b}$
- A **underdetermined** $\implies A^+\vec{b}$ gives least-squares solution to $A\vec{x} \approx \vec{b}$ with least (Euclidean) norm

Alternative Form

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

$$\ell \equiv \min\{m, n\}$$

Outer Product

$$\vec{u} \otimes \vec{v} \equiv \vec{u}\vec{v}^\top$$

Computing $A\vec{x}$

$$A\vec{x} = \sum_i \sigma_i(\vec{v}_i \cdot \vec{x})\vec{u}_i$$

Computing $A\vec{x}$

$$A\vec{x} = \sum_i \sigma_i (\vec{v}_i \cdot \vec{x})\vec{u}_i$$

Trick:

Ignore small σ_i.

Computing $A^+\vec{x}$

$$A^+ = \sum_{\sigma_i \neq 0} \frac{\vec{v}_i \vec{u}_i^\top}{\sigma_i}$$

Trick:

Ignore large σ_i.

Even Better Trick

Do not compute large (small) σ_i at all!

Eckart-Young Theorem

Theorem

Suppose $\tilde{A}$ is obtained from $A = U\Sigma V^\top$ by truncating all but the k largest singular values σ_i of A to zero. Then, $\tilde{A}$ minimizes both $\|A - \tilde{A}\|_{\text{Fro}}$ and $\|A - \tilde{A}\|_2$ subject to the constraint that the column space of $\tilde{A}$ has at most dimension k.

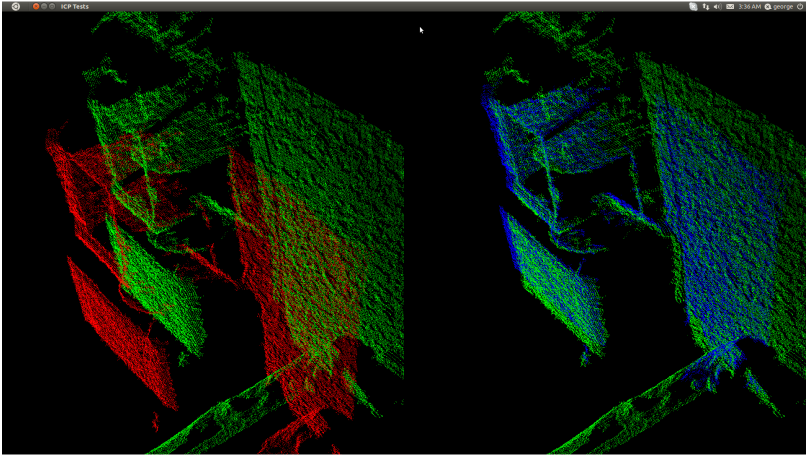
Matrix Norm Expressions

$$\|A\|_{\mathrm{Fro}}^2 = \sum_j \sigma_j^2$$

$$\|A\|_2 = \max\{\sigma_i\}$$

$$\operatorname{cond} A = \frac{\sigma_{\max}}{\sigma_{\min}}$$

Rigid Alignment



http://www.pointclouds.org/blog/_images/GICPRegistrationTopView.png

Procrustes Problem

Procrustes

From Wikipedia, the free encyclopedia

"Damastes" redirects here. For the spider genus, see Damastes (spider).

For the Larry Niven story, see Procrustes (short story).

In Greek mythology, **Procrustes** (Προκρούστης) or "the stretcher [who hammers out the metal]", also known as **Prokoptas** or **Damastes** (Δαμαστής) "subduer", was a rogue smith and bandit from Attica who physically attacked people by stretching them or cutting off their legs, so as to force them to fit the size of an iron bed. In general, when something is Procrustean, different lengths or sizes or properties are fitted to an arbitrary standard.

Contents [hide]

http://en.wikipedia.org/wiki/Procrustes

Procrustes via SVD

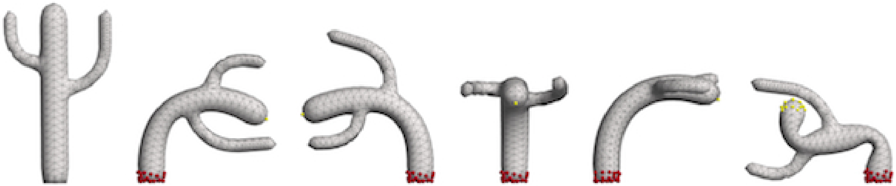
Orthogonal Procrustes Theorem

The orthogonal matrix R minimizing $\|RX - Y\|^2$ is given by $VU^\top$, where SVD is applied to factor $YX^\top = U\Sigma V^\top$.

Proved in notes.

Application: As-Rigid-As-Possible

As-Rigid-As-Possible Surface Modeling
Olga Sorkine and Marc Alexa
Eurographics/ACM SIGGRAPH Symposium on Geometry Processing 2007.
http://www.youtube.com/watch?v=1tX-qUjbkdc



Recall: Statistics Problem

Given: Collection of data points $\vec{x}_i$

- Age
- Weight
- Blood pressure
- Heart rate

Find: Correlations between different dimensions

Simplest Model

One-dimensional subspace

$$\vec{x}_i \approx c_i \vec{v}$$

More General Statement

Principal Components Analysis

The matrix $C \in \mathbb{R}^{n \times d}$ minimizing $\|X - CC^\top X\|_{\text{Fro}}$ subject to $C^\top C = I_{d \times d}$ is given by the first d columns of U, for $X = U\Sigma V^\top$.

Proved in notes.

Next