Model Order Reduction Techniques

SVD & POD

M. Grepl[a] & K. Veroy-Grepl[b]

[a]Institut für Geometrie und Praktische Mathematik

[b]Aachen Institute for Advanced Study in Computational Engineering Science (AICES)

RWTH Aachen

Sommersemester 2014

Singular Value Decomposition

Satz

For each matrix $\boldsymbol{Y} \in \mathbb{R}^{m \times n}$ there exists orthogonal matrices $\boldsymbol{U} \in \mathbb{R}^{m \times m}$, $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ and a diagonal matrix

$$\boldsymbol{\Sigma} := \operatorname{diag}(\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_p) \in \mathbb{R}^{m \times n}, \quad p = \min(m, n),$$

where

$$\boldsymbol{\sigma}_1 \geq \boldsymbol{\sigma}_2 \geq \ldots \geq \boldsymbol{\sigma}_p \geq 0,$$

such that

$$\boldsymbol{Y} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T.$$

Here:

- Singular values of $\boldsymbol{Y}$: $\boldsymbol{\sigma}_i,\ i = 1, \ldots, p$
- Left singular vectors : columns of $\boldsymbol{U} = [\boldsymbol{u}_1\, \boldsymbol{u}_2 \ldots \boldsymbol{u}_m]$
- Right singular vectors : columns of $\boldsymbol{V} = [\boldsymbol{v}_1\, \boldsymbol{v}_2 \ldots \boldsymbol{v}_n]$

Singular Value Decomposition

Case: $m = 4,\ n = 2$, $\mathbf{rank}(Y) = \mathbf{2}$

- Full SVD

$$\underbrace{\begin{bmatrix} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \times & 0 \\ 0 & \times \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}}_{V^T}$$

- Reduced SVD

$$\underbrace{\begin{bmatrix} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{bmatrix}}_{\hat{U}} \underbrace{\begin{bmatrix} \times & 0 \\ 0 & \times \end{bmatrix}}_{\hat{\Sigma}} \underbrace{\begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}}_{V^T}$$

Singular Value Decomposition

- Assumption: $m \geq n$
  - Case $m < n$: consider SVD of Y^T

$$Y^T = U\,\Sigma\,V^T \quad \Rightarrow \quad Y = V\,\Sigma\,U^T$$

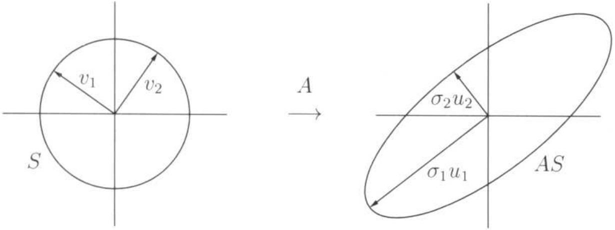
- Singular values are real and non-negative
- Convention
  - $\sigma_{\max} = \sigma_1$ largest singular value
  - $\sigma_{\min} = \sigma_n$ smallest singular values
  - ordered according to magnitude

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$$

- Numerical computation of SVD is backward stable

$$Y + \Delta Y = \tilde{U}\,\tilde{\Sigma}\,\tilde{V}^T \text{ with } \|\Delta Y\|_2 = \mathcal{O}(\text{eps})$$

Singular Value Decomposition – Geometry

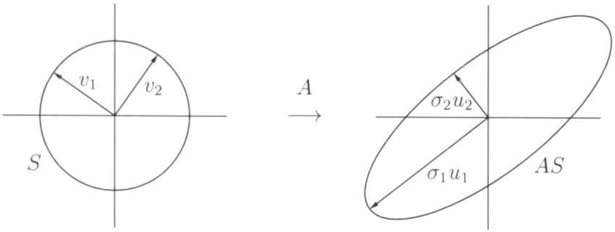


Quelle: Trefethen & Bau

Mapping of unit sphere $S = \{x : \|x\|_2 = 1\}$ under Y

- Singular values: lengths σ_1, σ_2 of principal semiaxes of YS.
- Left singular vectors: unit vectors $\{u_1, u_2\}$ in direction of principal semiaxes of YS
- Right singular vectors: unit vectors $\{v_1, v_2\} \in S$, preimages of principal semiaxes of YS, so that $Yv_j = \sigma_j u_j$.

Singular Value Decomposition – Geometry

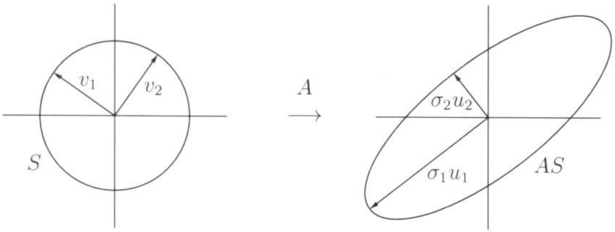


Quelle: Trefethen & Bau

Mapping of unit sphere $S = \{x : \|x\|_2 = 1\}$ under Y

- Singular values: lengths σ_1, σ_2 of principal semiaxes of YS.
- Left singular vectors: unit vectors $\{u_1, u_2\}$ in direction of principal semiaxes of YS
- Right singular vectors: unit vectors $\{v_1, v_2\} \in S$, preimages of principal semiaxes of YS, so that $Yv_j = \sigma_j u_j$.

Singular Value Decomposition – Geometry

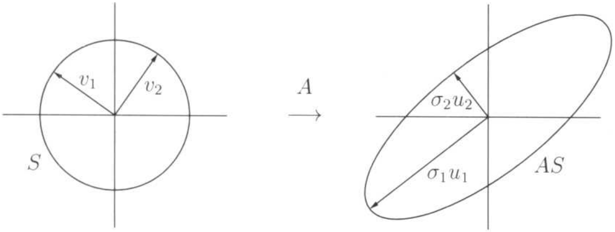


Quelle: Trefethen & Bau

Mapping of unit sphere $S = \{x : \|x\|_2 = 1\}$ under Y

- Singular values: lengths σ_1, σ_2 of principal semiaxes of YS.
- Left singular vectors: unit vectors $\{u_1, u_2\}$ in direction of principal semiaxes of YS
- Right singular vectors: unit vectors $\{v_1, v_2\} \in S$, preimages of principal semiaxes of YS, so that $Yv_j = \sigma_j u_j$.

Singular Value Decomposition – Geometry

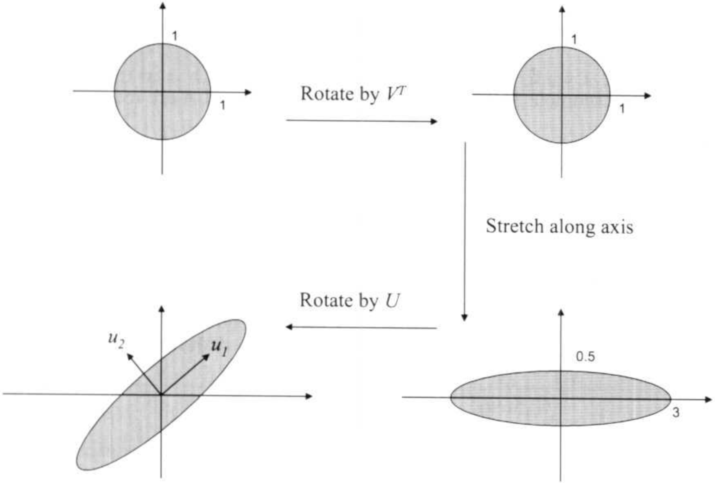


Quelle: Trefethen & Bau

Mapping of unit sphere $S = \{x : \|x\|_2 = 1\}$ under Y

- Singular values: lengths σ_1, σ_2 of principal semiaxes of YS.
- Left singular vectors: unit vectors $\{u_1, u_2\}$ in direction of principal semiaxes of YS
- Right singular vectors: unit vectors $\{v_1, v_2\} \in S$, preimages of principal semiaxes of YS, so that $Yv_j = \sigma_j u_j$.

Singular Value Decomposition – Geometry



Properties

Let $Y = U\,\Sigma\,V^T$ be an SVD of $Y \in \mathbb{R}^{m \times n}$ with singular values $\sigma_1 \geq \ldots \geq \sigma_d \geq \sigma_{d+1} = \ldots = \sigma_n = 0$. We then have:

- $Y v_i = \sigma_i u_i$, $Y^T u_i = \sigma_i v_i$, $i = 1, \ldots, n$.
- $\mathrm{Rang}(Y) = d =$ number of singular values not equal to zero.
- $\|Y\|_2 = \sigma_1 = \sigma_{\max}$
- $\|Y\|_{\mathrm{F}} = \sqrt{\sigma_1^2 + \ldots + \sigma_n^2}$
- $\|Y^{-1}\|_2 = \frac{1}{\sigma_n}$, if $Y \in \mathbb{R}^{n \times n}$ regular
- $\kappa_2(Y) = \|Y\|_2 \|Y^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}$, if $Y \in \mathbb{R}^{n \times n}$ regular
- $\kappa_2(Y) = \|Y\|_2 \|Y^{+}\|_2 = \frac{\sigma_1}{\sigma_n} = \frac{\sigma_{\max}}{\sigma_{\min}}$, if Y regular

Properties

Let $\boldsymbol{Y} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T$ be an SVD of $\boldsymbol{Y} \in \boldsymbol{R}^{m\times n}$ with singular values $\boldsymbol{\sigma}_1 \geq \ldots \geq \boldsymbol{\sigma}_d \geq \boldsymbol{\sigma}_{d+1} = \ldots = \boldsymbol{\sigma}_n = \boldsymbol{0}$. We then have:

- The strictly positiv singular values are the roots of the strictly positiv eigenvalues of $\boldsymbol{Y}^T\boldsymbol{Y}$:
  $$\{\boldsymbol{\sigma}_i \,|\, i = 1, \ldots, d\} = \{\sqrt{\boldsymbol{\lambda}_i(\boldsymbol{Y}^T\boldsymbol{Y})} \,|\, i = 1, \ldots, d\}$$
- The singular values are equal to the absolute values of the eigenvalues of $\boldsymbol{Y}$ if $\boldsymbol{Y} = \boldsymbol{Y}^T$.
- For $\boldsymbol{Y} \in \mathbb{R}^{n\times n}$ it holds $|\det(\boldsymbol{Y})| = \prod_{i=1}^{m} \boldsymbol{\sigma}_i$
- The pseudo inverse $\boldsymbol{Y}^+ \in \mathbb{R}^{n\times m}$ is defined as
  $$\boldsymbol{Y}^+ = \boldsymbol{V}\,\boldsymbol{\Sigma}^+\,\boldsymbol{U}^T$$
  where $\boldsymbol{\Sigma}^+ = \mathbf{diag}(\boldsymbol{\sigma}_1^{-1}, \ldots, \boldsymbol{\sigma}_d^{-1}, 0, \ldots, 0) \in \mathbb{R}^{n\times m}$.

Properties – Low-Rank Approximation

Let $Y = U\,\Sigma\,V^T$ be an SVD of $Y \in R^{m\times n}$ with singular values $\sigma_1 \geq \ldots \geq \sigma_d \geq \sigma_{d+1} = \ldots = \sigma_n = 0$. For $k \leq d = \mathrm{Rang}(Y)$ define

$$Y_k = U\,\Sigma_k\,V^T$$

where $\Sigma_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0) \in \mathbb{R}^{m\times n}$. We then have:

- $\mathrm{Rang}(Y) = k$.
- The distance between Y_k and Y in the 2-norm is
$$\|Y - Y_k\|_2 = \sigma_{k+1}.$$
- Y_k is the best approximation of Y of rank $\leq k$
$$\|Y - Y_k\|_2 = \min_{\mathrm{Rang}(B)\leq k} \|Y - B\|_2.$$

Image Compression

Singular values decomposition of Y

$$Y = U\,\Sigma\,V^T$$

can be written as

$$Y = \sum_{i=1}^{p} \sigma_i\, u_i\, v_i^T = \sigma_1\, u_1\, v_1^T + \sigma_2\, u_2\, v_2^T + \ldots + \sigma_p\, u_p\, v_p^T$$

where $p = \min(m, n)$.

Approximation of matrix Y of rank k, $k \leq p$, is

$$Y_k = \sum_{i=1}^{k} \sigma_i\, u_i\, v_i^T$$

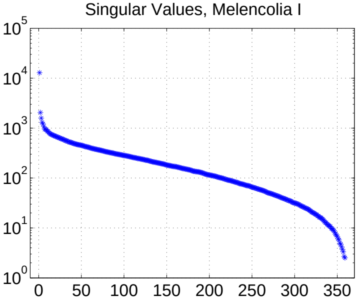
and $\|Y - Y_k\|_2 = \sigma_{k+1}$.

Data compression:

$\Rightarrow$ Memory requirement of Y_k is $k(m+n)$ vs. mn for Y.

Image Compression

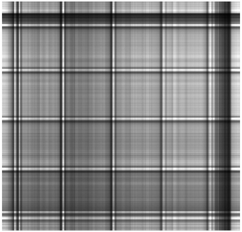
Consider black and white picture $I(x, y)$ as a matrix Y: matrix entries correspond to the gray-level of the pixels.



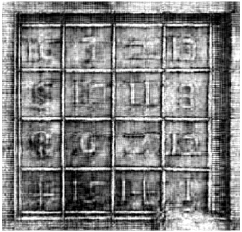
Depending on the area of application: Principal Component Analysis (PCA), Proper Orthogonal Decomposition (POD), Karhunen-Loéve Decomposition.

Image Compression

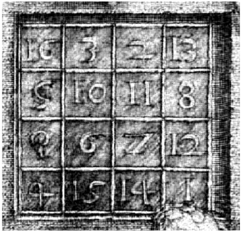
k = 1, compression = 0.00548



k = 20, compression = 0.11



k = 40, compression = 0.219



Original: A. Duerer, Melencolia I

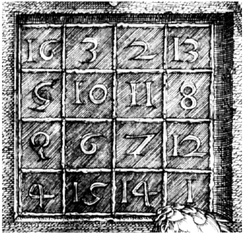


Image Compression

Gatlinburg Conference: J.H. Wilkinson, W. Givens, G. Forsythe, A. Householder, G. Henrici, F.L. Bauer (von links nach rechts)

SVD & POD

We consider the problem of approximating all spatial vectors $\{y_j\}_{j=1}^n$ of Y simultaneously by a single, normalized vector as well as possible. This problem can be expressed as

$$(\mathrm{P}^1) \quad \max_{u \in \mathbb{R}^m} \sum_{j=1}^{n} |(y_j, u)_{\mathbb{R}^m}|^2 \quad \text{subject to (s.t.)} \quad \|u\|_{\mathbb{R}^m}^2 = 1$$

Solution: u_1 solves (P^1) and $\arg\max(\mathrm{P}^1) = \sigma_1^2 = \lambda_1$.

Find second vector, orthogonal to u_1, that describes $\{y_j\}_{j=1}^n$ as well as possible

$$(\mathrm{P}^2) \quad \max_{u \in \mathbb{R}^m} \sum_{j=1}^{n} |(y_j, u)_{\mathbb{R}^m}|^2$$

$$\text{s.t. } \|u\|_{\mathbb{R}^m}^2 = 1, \ (u, u_1)_{\mathbb{R}^m} = 0$$

Solution: u_2 solves (P^2) and $\arg\max(\mathrm{P}^2) = \sigma_2^2 = \lambda_2$.

SVD & POD

We consider the problem of approximating all spatial vectors $\{y_j\}_{j=1}^n$ of Y simultaneously by a single, normalized vector as well as possible. This problem can be expressed as

$$(\mathrm{P}^1) \quad \max_{u \in \mathbb{R}^m} \sum_{j=1}^{n} |(y_j, u)_{\mathbb{R}^m}|^2 \quad \text{subject to (s.t.)} \quad \|u\|_{\mathbb{R}^m}^2 = 1$$

Solution: u_1 solves (P^1) and $\arg\max(\mathrm{P}^1) = \sigma_1^2 = \lambda_1$.

Find second vector, orthogonal to u_1, that describes $\{y_j\}_{j=1}^n$ as well as possible

$$(\mathrm{P}^2) \quad \max_{u \in \mathbb{R}^m} \sum_{j=1}^{n} |(y_j, u)_{\mathbb{R}^m}|^2$$

$$\text{s.t. } \|u\|_{\mathbb{R}^m}^2 = 1, \; (u, u_1)_{\mathbb{R}^m} = 0$$

Solution: u_2 solves (P^2) and $\arg\max(\mathrm{P}^2) = \sigma_2^2 = \lambda_2$.

SVD & POD

Theorem

Let $Y = [y_1, \dots, y_n] \in \mathbb{R}^{m\times n}$ be a given matrix with rank $d \leq \min\{m, n\}$. Further, let $Y = U\Sigma V^T$ be the singular value decomposition of Y, where $U = [u_1, \dots, u_m] \in \mathbb{R}^{m\times m}$, $V = [v_1, \dots, v_n] \in \mathbb{R}^{n\times n}$ are orthogonal matrices and the matrix $\Sigma \in \mathbb{R}^{m\times n}$ contains the singular values. Then, for any $\ell \in \{1, \dots, d\}$ the solution to

$$(\mathrm{P}^\ell) \qquad \max_{\tilde{u}_1,\dots,\tilde{u}_\ell \in \mathbb{R}^m} \sum_{i=1}^{\ell} \sum_{j=1}^{n} |(y_j, \tilde{u}_i)_{\mathbb{R}^m}|^2$$

$$\text{s.t.} \quad (\tilde{u}_i, \tilde{u}_j)_{\mathbb{R}^m} = \delta_{ij}, \; 1 \leq i, j \leq \ell$$

is given by the singular vector $\{u_i\}_{i=1}^{\ell}$, i.e., by the first ℓ columns of U. Moreover,

$$\arg\max(\mathrm{P}^\ell) = \sum_{i=1}^{\ell} \sigma_i^2 = \sum_{i=1}^{\ell} \lambda_i.$$

Optimality of the POD basis

Theorem

Suppose that $\hat{U}^d \in \mathbb{R}^{m \times d}$ denotes a matrix with pairwise orthonormal vectors $\hat{u}_i$ and that the expansion of the columns of Y in the basis $\{\hat{u}_i\}_{i=1}^d$ be given by

$$Y = \hat{U}^d C^d, \text{ with } C_{ij}^d = (\hat{u}_i, y_j)_{\mathbb{R}^m}, \; 1 \leq i \leq d, 1 \leq j \leq n.$$

Then for every $\ell \in \{1, \ldots, d\}$ we have

$$\|Y - U^\ell B^\ell\|_F \leq \|Y - \hat{U}^\ell C^\ell\|_F$$

where $\|\cdot\|_F$ denotes the Frobenius norm, the matrix U^ℓ denotes the first ℓ columns of U, B^ℓ the first ℓ rows of B and similarly for $\hat{U}^\ell$ and C^ℓ.

S. Volkwein. Model Reduction Using Proper Orthogonal Decomposition. *Skript Universität Konstanz*, 2010.

A POD Approach

Proper Orthogonal Decomposition (POD) (or Karhunen-Loève expansion) approach:

- POD Spaces

$$X_N^{\text{POD}} = \arg \inf_{\text{spaces } X_N \subset \text{span}\{u(\mu) | \mu \in \Xi_{\text{train}}\}} \|u - \Pi_{X_N} u\|_{L^2(\Xi_{\text{train}};X)}$$

- "Best" approximation error

$$\bar{\bar{\varepsilon}}_N^{\text{POD}} \equiv \|u - \Pi_{X_N^{\text{POD}}} u\|_{L^2(\Xi_{\text{train}};X)}$$

Note:

- X_N^{POD} are hierarchical,
- Weaker norm over Ξ_{train},
- Optimization can be solved using "method of snapshosts."

A POD Approach – Method of Snapshots

1. Form correlation matrix $\underline{C}^{\mathrm{POD}} \in \mathbb{R}^{n_{\mathrm{train}} \times n_{\mathrm{train}}}$ given by
$$C_{ij}^{\mathrm{POD}} = \frac{1}{n_{\mathrm{train}}} (u(\mu_{\mathrm{train}}^i), u(\mu_{\mathrm{train}}^j))_X, \quad 1 \leq i, j \leq n_{\mathrm{train}}.$$

2. Solve for eigenpairs $(\underline{\psi}^{\mathrm{POD},k} \in \mathbb{R}^{n_{\mathrm{train}}}, \lambda^{\mathrm{POD},k} \in \mathbb{R}_{+0})$, $1 \leq k \leq n_{\mathrm{train}}$, from
$$\underline{C}^{\mathrm{POD}} \, \underline{\psi}^{\mathrm{POD},k} = \lambda^{\mathrm{POD},k} \underline{\psi}^{\mathrm{POD},k},$$
$$(\underline{\psi}^{\mathrm{POD},k})^T \underline{\mathbb{X}} \, \underline{\psi}^{\mathrm{POD},k} = 1.$$

3. Arrange eigenvalues in descending order
$$\lambda^{\mathrm{POD},1} \geq \lambda^{\mathrm{POD},1} \geq \ldots \geq \lambda^{\mathrm{POD},n_{\mathrm{train}}} \geq 0.$$

A POD Approach – Method of Snapshots

4. POD Basis functions

$$\Psi^{\text{POD},k} \equiv \sum_{m=1}^{n_{\text{train}}} \psi_m^{\text{POD},k} u(\mu_{\text{train}}^m), \quad 1 \le k \le n_{\text{train}}.$$

$\Rightarrow$ POD reduced basis space

$$X_N^{\text{POD}} = \text{span}\{\Psi^{\text{POD},n},\ 1 \le n \le N\}, \quad 1 \le N \le N_{\max}$$

$\Rightarrow$ Error bound: define $N_{\max}$ as the smallest N such that

$$(\bar{\bar{\varepsilon}}_N^{\text{POD}} \equiv) \sqrt{\sum_{k=N+1}^{n_{\text{train}}} \lambda^{\text{POD},k}} \le \varepsilon_{\text{tol,min}}.$$

A POD Approach – Method of Snapshots

Since $(\boldsymbol{\Psi}^{\mathbf{POD},n}, \boldsymbol{\Psi}^{\mathbf{POD},n})_X = \delta_{nm},\ 1 \le n, m \le n_{\mathrm{train}}$, it follows that

$$(\boldsymbol{\Psi}^{\mathbf{POD},n} \equiv)\xi^n = \zeta^n, \quad 1 \le n \le N_{\max},$$

and hence $\underline{\mathbb{Z}} \in \mathbb{R}^{\mathcal{N} \times N}$ is given by

$$\underline{\mathbb{Z}} = [\boldsymbol{\Psi}^{\mathbf{POD},1}\ \boldsymbol{\Psi}^{\mathbf{POD},2} \ldots \boldsymbol{\Psi}^{\mathbf{POD},N}].$$

Note:

- n_{train} $\underline{A}^{\mathcal{N}}$-solves and n^2_{train} X-inner products to form $\underline{C}$
- Solve for largest $N_{\max}$ eigenvalues/eigenvectors of $\underline{C}$

$\Rightarrow$ higher parameter dimensions prohibitive because of large n_{train}.