

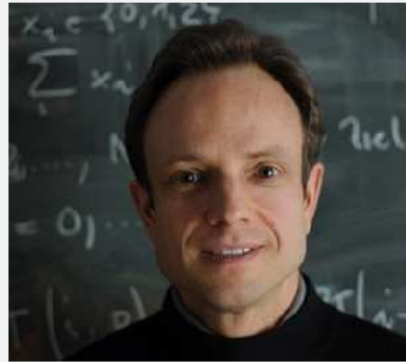
Course Video

Section 3.4.4.2: Best Low-Rank Approximation

Prof. R. Hiptmair, SAM, ETH Zurich

Date: September 29, 2020

(C) Seminar für Angewandte Mathematik, ETH Zürich



Prerequisites.

- Matrix norms

Dependencies. Section 3.3.2, Section 3.4.1, Section 3.4.2

Duration: 25 minutes

III. Direct Methods for Linear Least Squares Problems

3.4. Singular Value Decomposition (SVD)

3.4.4. SVD-Based Optimization and Approximation

3.4.4.2. Best Low-Rank Approximation

Context: Matrix compression = approximation by *data-sparse* matrix

← sparse matrices → low-rank matrices

$$r := \text{rank}(A) \ll \min\{m, n\} \\ \text{for } A \in \mathbb{R}^{m,n}$$

If $\text{rank}(A) = r$:

[by SVD] $\Rightarrow A = \sum_{\ell=1}^r \sigma_{\ell} (U)_{:, \ell} (V)_{:, \ell}^T$

$$\text{Storage} \approx r(m+n) \ll m \cdot n$$

Theorem 3.4.4.18. Best low rank approximation →

Let $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ be the SVD of $\mathbf{A} \in \mathbb{K}^{m,n}$ ($\rightarrow$ Thm. 3.4.1.1). For $1 \leq k \leq \text{rank}(\mathbf{A})$ set

$$\mathbf{A}_k := \underbrace{\mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H}_{\text{truncated SVD}} = \sum_{\ell=1}^k \sigma_\ell (\mathbf{U})_{:, \ell} (\mathbf{V})_{:, \ell}^H \quad \text{with} \quad \begin{aligned} \mathbf{U}_k &:= \begin{bmatrix} (\mathbf{U})_{:,1}, \dots, (\mathbf{U})_{:,k} \end{bmatrix} \in \mathbb{K}^{m,k}, \\ \mathbf{V}_k &:= \begin{bmatrix} (\mathbf{V})_{:,1}, \dots, (\mathbf{V})_{:,k} \end{bmatrix} \in \mathbb{K}^{n,k}, \\ \mathbf{\Sigma}_k &:= \text{diag}(\sigma_1, \dots, \sigma_k) \in \mathbb{K}^{k,k}. \end{aligned}$$

Then, for $\|\cdot\| = \|\cdot\|_F$ and $\|\cdot\| = \|\cdot\|_2$, holds true

$$\|\mathbf{A} - \mathbf{A}_k\| \leq \|\mathbf{A} - \mathbf{F}\| \quad \forall \mathbf{F} \in \mathcal{R}_k(m, n),$$

that is, $\mathbf{A}_k$ is the rank- k best approximation of $\mathbf{A}$ in the matrix norms $\|\cdot\|_F$ and $\|\cdot\|_2$.

$$\mathcal{R}_k(m, n) = \{ \mathbf{M} \in \mathbb{K}^{m,n} : \text{rank}(\mathbf{M}) \leq k \}$$

Definition 3.4.4.16. Frobenius norm

The **Frobenius norm** of $\mathbf{A} \in \mathbb{K}^{m,n}$ is defined as

$$\|\mathbf{A}\|_F^2 := \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2. \quad \triangleq \text{Euclidean vector norm for matrices}$$

Approximation error norm: ($m \geq n$)

$$\begin{aligned} \text{Full SVD: } \mathbf{A} &= \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma}_k & 0 \\ 0 & \sigma_{k+1} \dots \sigma_n \\ 0 & 0 \end{bmatrix} \mathbf{V}^H \\ \mathbf{A}_k &= \mathbf{U} \begin{bmatrix} \mathbf{\Sigma}_k & 0 \\ 0 & 0 \end{bmatrix} \mathbf{V}^H \end{aligned}$$

$$\|\mathbf{A} - \mathbf{A}_k\|_{2/F} = \left\| \mathbf{U} \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{k+1} \dots \sigma_n \\ 0 & 0 \end{bmatrix} \mathbf{V}^H \right\|_{2/F} = \sigma_{k+1}$$

$\|\cdot\|_F = \sigma_{k+1}^2 + \dots + \sigma_n^2$

because $\|\cdot\|_{2/F}$ is invariant w.r.t. multiplication with orthogonal matrices.

$$\text{In Euclidean matrix norm: } \|\mathbf{A} - \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H\|_2 = \sigma_{k+1}, \quad (3.4.4.21)$$

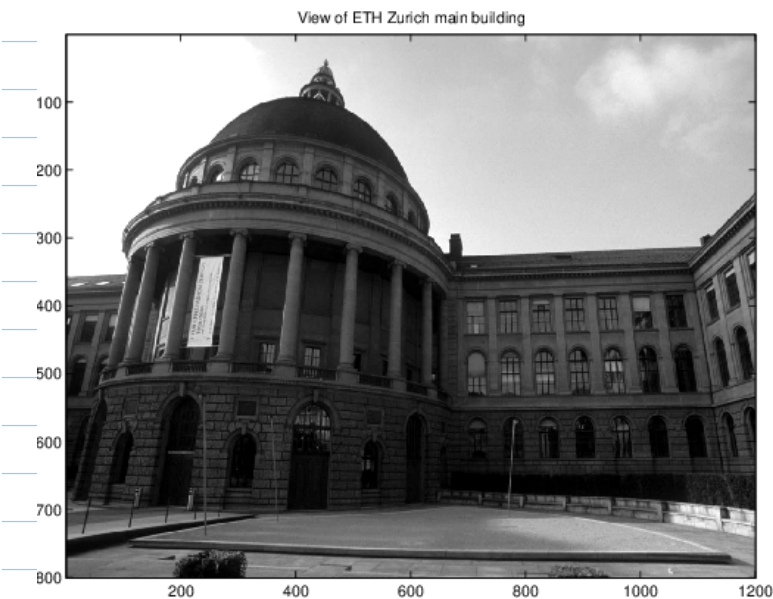
$$\text{in Frobenius norm: } \|\mathbf{A} - \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^H\|_F^2 = \sum_{j=k+1}^{\min\{m,n\}} \sigma_j^2. \quad (3.4.4.22)$$

▷ Decay of SV predicts quality of low-rank approximation

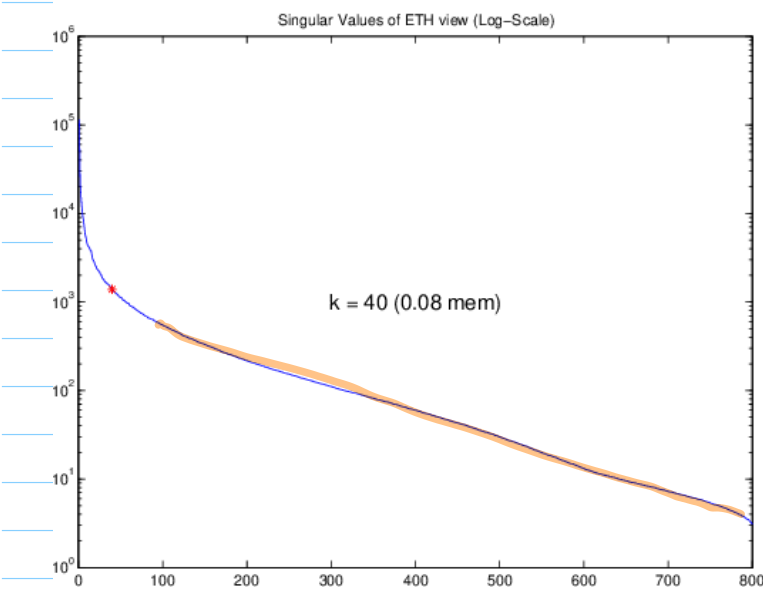
$$\text{Relative compression error } \frac{\|\mathbf{A} - \mathbf{A}_k\|_2}{\|\mathbf{A}\|_2} = \frac{\sigma_{k+1}}{\sigma_1} \stackrel{!}{\leq} \text{tol}$$

③

Ex 3.4.4.23 SVD for image compression



BMP grayscale image
 $\leftrightarrow$ matrix $A \in \{0, \dots, 255\}^{800, 1200}$



△ lin-log scale

! Fast (exponential) decay of the singular values

$$\sigma_i \sim q^i$$

for $0 < q < 1$

Review questions

Please answer review questions 3.4.4.25

4

5

⑦