Medical Image Processing for Interventional Applications

SVD in Optimization - Part 1

Online Course – Unit 4

Andreas Maier, Joachim Hornegger, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Optimization Problem I

Let us consider the following problem that appears in many image processing and computer vision problems:

We computed a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ out of sensor data, like an image.
By theory the matrix $\boldsymbol{A}$ must have the singular values

$$\sigma_1, \sigma_2, \ldots, \sigma_k, \; k \leq p = \min\{m, n\}.$$

Of course, in practice $\boldsymbol{A}$ does not always satisfy this constraint.

Problem: What is the matrix $\boldsymbol{A}' \in \mathbb{R}^{m \times n}$ that is closest to $\boldsymbol{A}$ (according to the Frobenius norm) and has the required singular values?

Solution: Let $\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^{\mathsf{T}}$, then

$$\boldsymbol{A}' = \boldsymbol{U} \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k) \boldsymbol{V}^{\mathsf{T}}.$$

Optimization Problem I

Example

Our measurements lead to the following matrix:

$$\boldsymbol{A} = \begin{pmatrix} 11 & 10 & 14 \\ 12 & 11 & -13 \\ 14 & 13 & -66 \end{pmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}}.$$

Let us assume that by theoretical arguments the matrix $\boldsymbol{A}$ is required to have a rank deficiency of one, and the two non-zero singular values are identical. The matrix $\boldsymbol{A}'$ that is closest to $\boldsymbol{A}$ w. r. t. the Frobenius norm and fulfills the requirements above is:

$$\boldsymbol{A}' = \boldsymbol{U}\,\mathrm{diag}\left(\frac{71.3967+21.7831}{2}, \frac{71.3967+21.7831}{2}, 0\right)\boldsymbol{V}^{\mathsf{T}}.$$

Optimization Problem II

Problem: In image processing we are often required to solve the following optimization problem:

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \boldsymbol{x}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\boldsymbol{x}, \quad \text{subject to} \quad \|\boldsymbol{x}\|_2 = 1,$$

or in the extreme:

$$\boldsymbol{A}\boldsymbol{x} = 0, \quad \text{subject to} \quad \|\boldsymbol{x}\|_2 = 1.$$

Solution: The solution can be constructed using the rightmost column of $\boldsymbol{V}$.

Exercise: Check this!

Optimization Problem II

Example

Estimate the matrix $\boldsymbol{X} \in \mathbb{R}^{2\times 2}$ such that for vectors

$$\boldsymbol{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{b}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \boldsymbol{b}_3 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \boldsymbol{b}_4 = \begin{pmatrix} -1 \\ -4 \end{pmatrix},$$

the following optimization problem gets solved:

$$\begin{aligned} & \sum_{i=1}^{4} \left(\boldsymbol{b}_i^\top \boldsymbol{X} \boldsymbol{b}_i\right)^2 \rightarrow \min, \quad \text{subject to} \quad \|\boldsymbol{X}\|_F = 1, \\ \Leftrightarrow \quad & \boldsymbol{b}_i^\top \boldsymbol{X} \boldsymbol{b}_i = 0, \quad i = 1, \ldots, 4, \quad \|\boldsymbol{X}\|_F = 1. \end{aligned}$$

Optimization Problem II

Example

The objective function is linear in the components of $\boldsymbol{X}$, thus the whole sum can be rewritten in matrix notation:

$$\boldsymbol{M}\boldsymbol{x} = \boldsymbol{M}\begin{pmatrix} x_{1,1} \\ x_{1,2} \\ x_{2,1} \\ x_{2,2} \end{pmatrix} = 0, \quad \text{subject to} \quad \|\boldsymbol{x}\|_2 = 1,$$

where the ***measurement matrix*** $\boldsymbol{M}$ is built from single elements of the sum.

Optimization Problem II

Example

Let us consider the i-th component:

$$\boldsymbol{b}_i^{\mathrm{T}} \boldsymbol{X} \boldsymbol{b}_i = \boldsymbol{b}_i^{\mathrm{T}} \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} \boldsymbol{b}_i = \left(b_{i,1}^2, b_{i,1}b_{i,2}, b_{i,1}b_{i,2}, b_{i,2}^2\right) \begin{pmatrix} x_{1,1} \\ x_{1,2} \\ x_{2,1} \\ x_{2,2} \end{pmatrix}.$$

Using this result we get for the overall measurement matrix:

$$\boldsymbol{M} = \begin{pmatrix} b_{1,1}^2 & b_{1,1}b_{1,2} & b_{1,1}b_{1,2} & b_{1,2}^2 \\ b_{2,1}^2 & b_{2,1}b_{2,2} & b_{2,1}b_{2,2} & b_{2,2}^2 \\ b_{3,1}^2 & b_{3,1}b_{3,2} & b_{3,1}b_{3,2} & b_{3,2}^2 \\ b_{4,1}^2 & b_{4,1}b_{4,2} & b_{4,1}b_{4,2} & b_{4,2}^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -2 & -2 & 4 \\ 1 & -3 & -3 & 9 \\ 1 & 4 & 4 & 16 \end{pmatrix}$$

Optimization Problem II

Example

The nullspace of $\boldsymbol{M}$ can be computed by SVD and yields the required matrix:

$$\boldsymbol{X} = \begin{pmatrix} 0 & -0.7071 \\ 0.7071 & 0 \end{pmatrix},$$

which satisfies $\|\boldsymbol{X}\|_F \approx 1$.

Take Home Messages

- For a theoretically known rank, we can compute a best approximation of a matrix computed from measurements using SVD.
- SVD can be applied to many optimization problems (see also next unit).
- When it comes to optimization, always check if you can solve it with SVD.

Further Readings

Read the original:

Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. 3rd ed. Johns Hopkins Studies in the Mathematical Sciences. Baltimore: The Johns Hopkins University Press, Oct. 1996

A very detailed and easy to follow introduction of the SVD can be found in:

Carlo Tomasi's class notes, chapter 3 (a **must-read**).

The theory is described in an easy to read format in:

Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*. Philadelphia: SIAM, 1997

For details about the numerical computation of SVD see:

William H. Press et al. *Numerical Recipes – The Art of Scientific Computing*. 3rd ed. Cambridge University Press, 2007. Get at `http://numerical.recipes/` (August 2016).

A good reference for properties of matrices is the following script:

Kaare Brandt Petersen and Michael Syskind Pedersen. *The Matrix Cookbook*. Online. Technical University of Denmark, Nov. 2012. URL: `http://www2.imm.dtu.dk/pubdb/p.php?3274`