Medical Image Processing for Interventional Applications

Teaser – Course Introduction

Online Course – Unit 1

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

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Overview

- Image preprocessing and image enhancement
- Image information extraction
- Non-rigid image registration
- Interventional reconstruction

Mathematical Tools

The lecture will emphasize on various mathematical tools:

- linear algebra,
- discrete differentiation,
- local and global optimization,
- variational calculus,
- partial differential equations.

Image Preprocessing

- Edge detection
- Hough transform
- Structure tensor
- Vesselness filter

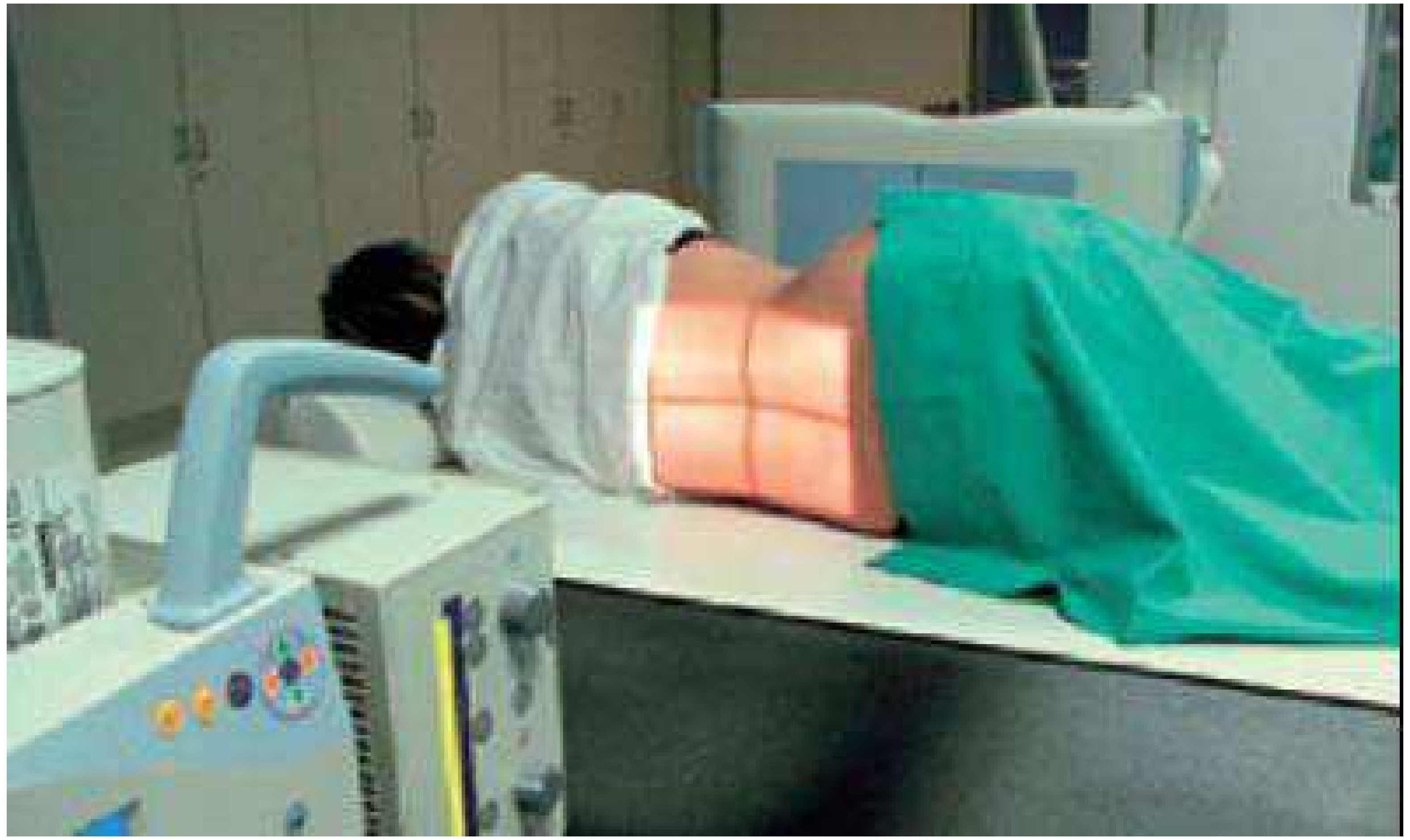
Image Enhancement

- Linear filtering
- Joint bilateral filtering
- Guided filter
- Super resolution

Image Information Extraction

- Epipolar geometry and consistency
- Structure from motion (factorization)
- Localization of organs
- Segmentation of cysts in kidneys using random walk algorithm
- Active shape models

Shutter Correction



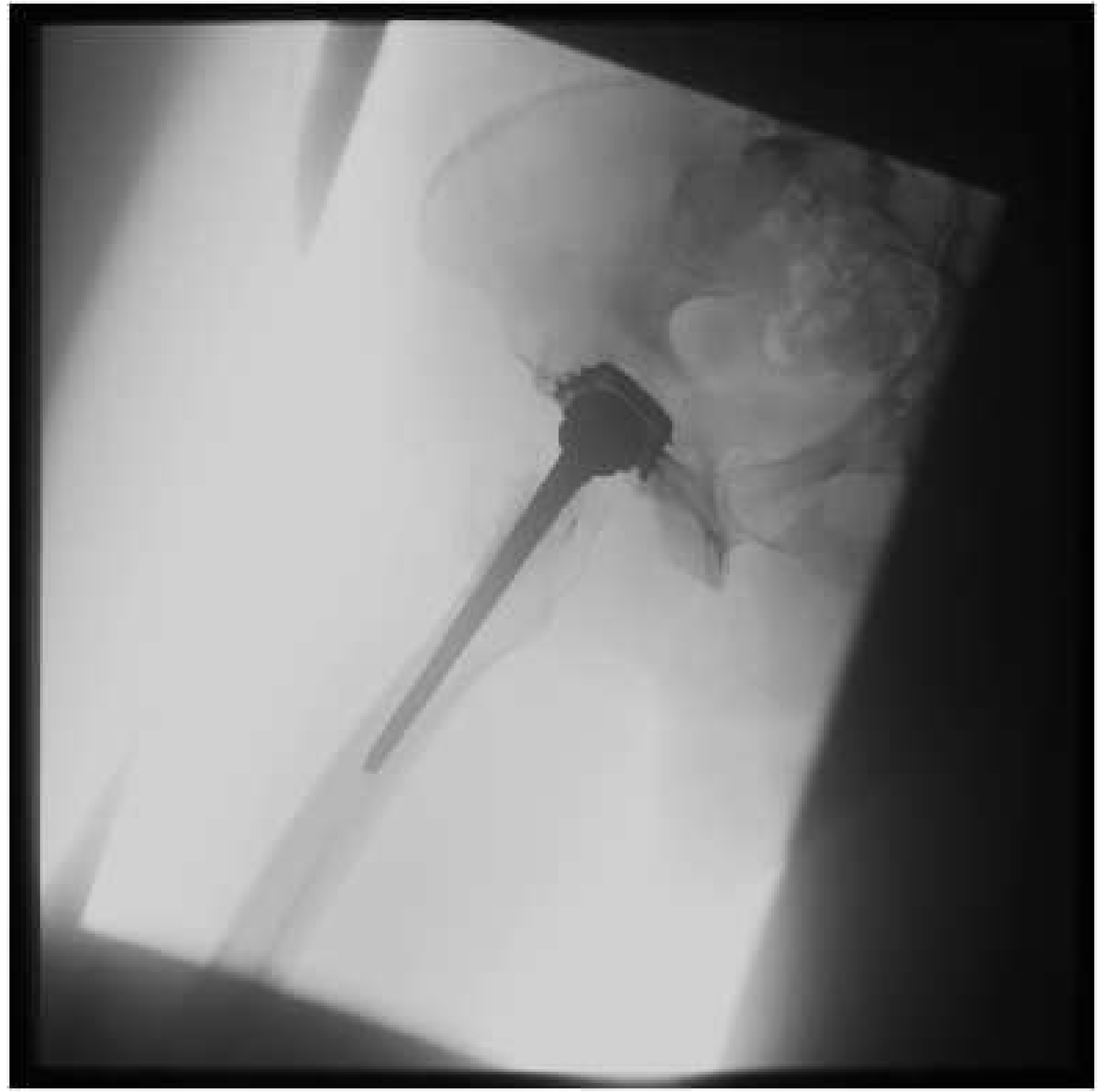


Figure 1: Shutter – detection of boundaries required

Example Images: Shutter Detection

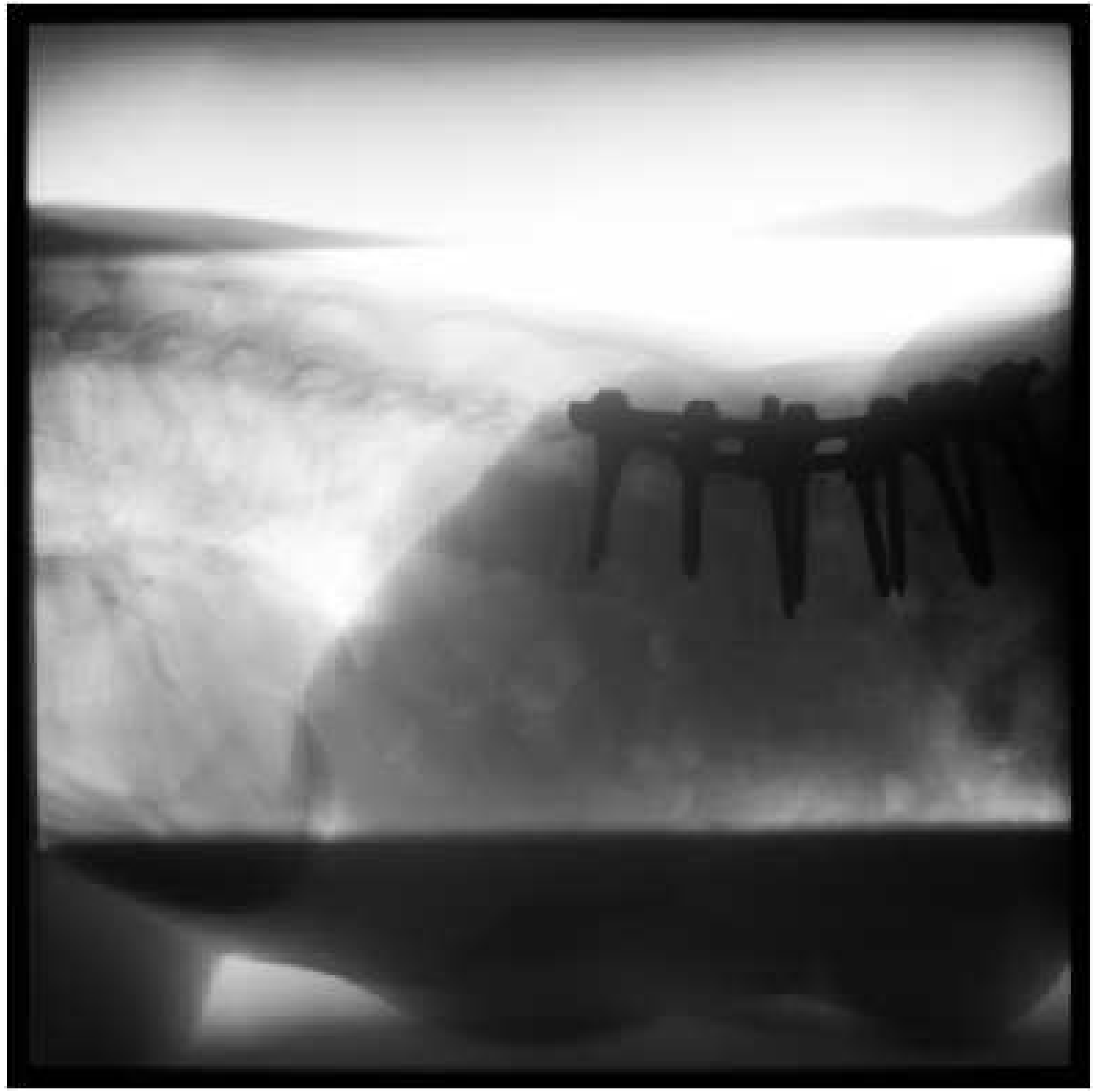


Figure 2: Shutter – detection of boundaries required

Example Images: Shutter Detection

Figure 3: Example for shutter detection (Klaus Finkler, Siemens Healthcare)

Computer Guided Intervention and Surgery

- Magnetic catheter navigation
- Augmented reality in minimal invasive surgery
- NOTES: **N**on **o**bstructive **t**ransluminal **e**ndoscopic **s**urgery

Figure 4: Magnetic navigation (image courtesy of Siemens Healthcare)

Computer Guided Intervention and Surgery

Figure 5: Brainlab navigation system

Non-Rigid Image Registration

- Digital subtraction imaging
- Non-rigid mono-modal image registration (variational approach)
- Non-rigid image registration using geometric constraints
- Non-rigid image registration using prior knowledge

Interventional Reconstruction

- 3-D ultrasound reconstruction
- Motion compensated reconstruction:
  - ECG gating
  - Motion compensated reconstruction using deformation
  - ECG guided motion detection

What makes Medical Image Processing for Interventional Applications so special?

- Hands-on-hardware lecture
- Insight into research results from important scientific publications
- Topics of high interest for industry projects
- Excellent preparation for bachelor's or master's thesis projects
- **Difficult, but tons of fun!**

Medical Image Processing for Interventional Applications

Singular Value Decomposition

Online Course – Unit 2
Andreas Maier, Joachim Hornegger, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Singular Value Decomposition

- Powerful normal form for matrices that allows for a simple solution of most linear problems in imaging and image processing.
- SVD is a method from linear algebra ...
  - ... invented in the 19th century.
  - ... rediscovered and pushed for practical applications by Gene Golub.
  - ... established in computer vision by Carlo Tomasi's famous factorization algorithm to compute structure and camera motion from image sequences.
  - ... which is extremely robust and simple to use.

Singular Value Decomposition

SVD is a perfect tool, e. g., for

- the computation of singular values,
- the computation of the null space,
- the computation of the (pseudo-) inverse,
- the solution of overdetermined linear equations,
- the computation of condition numbers,
- enforcing a rank criterion (numerical rank),
- and other applications of matrices.

On the Geometry of Linear Mappings

From linear algebra, we know that a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ maps the unit vectors $\boldsymbol{e}_i \in \mathbb{R}^n$ of the standard base to the corresponding column vectors $\boldsymbol{a}_i \in \mathbb{R}^m$ of the matrix $\boldsymbol{A}$, $i = 1, \ldots, n$.

Example

$$\boldsymbol{A} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_6) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{a}_4$$

On the Geometry of Linear Mappings

Example

Compute the orthogonal matrix $\boldsymbol{R}$, i. e., $\boldsymbol{R}^{-1} = \boldsymbol{R}^{\mathsf{T}}$, that rotates points in the 2-D image plane by the angle θ.

Solution:
The base vectors are mapped as follows:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix},$$

and thus the 2-D rotation matrix is:

$$\boldsymbol{R} \quad = \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

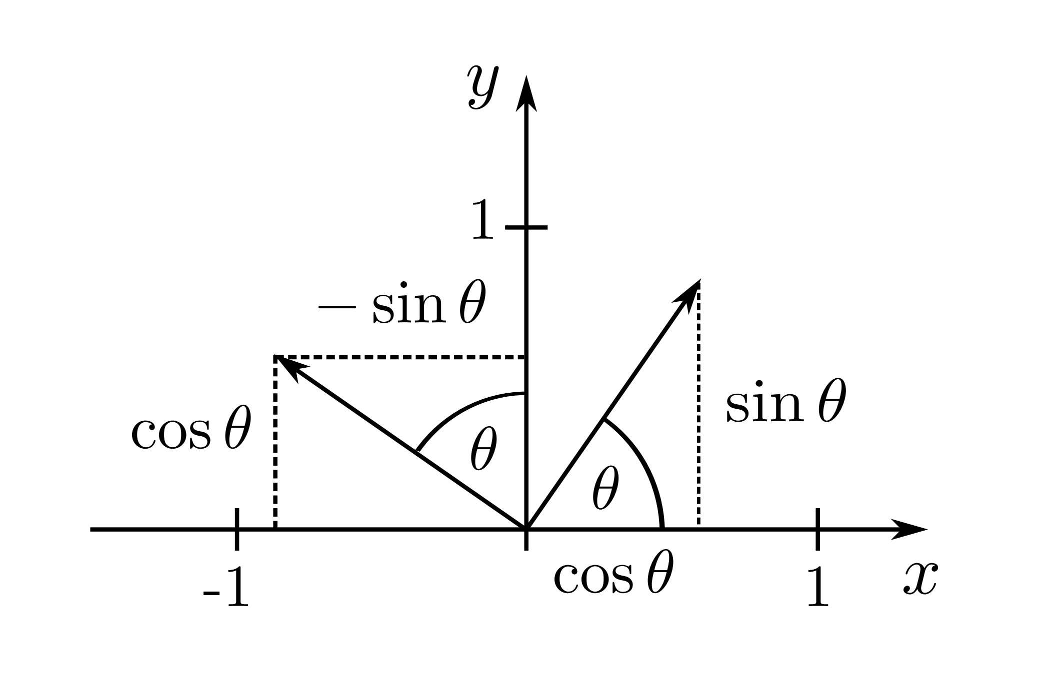


Figure 1: Rotation of 2-D unit vectors

On the Geometry of Linear Mappings

If $\boldsymbol{A}$ is a real $m \times n-$matrix of rank r, then $\boldsymbol{A}$ maps the unit hyper-sphere in the n-dimensional space to an r-dimensional hyperellipsoid in the m-dimensional space.

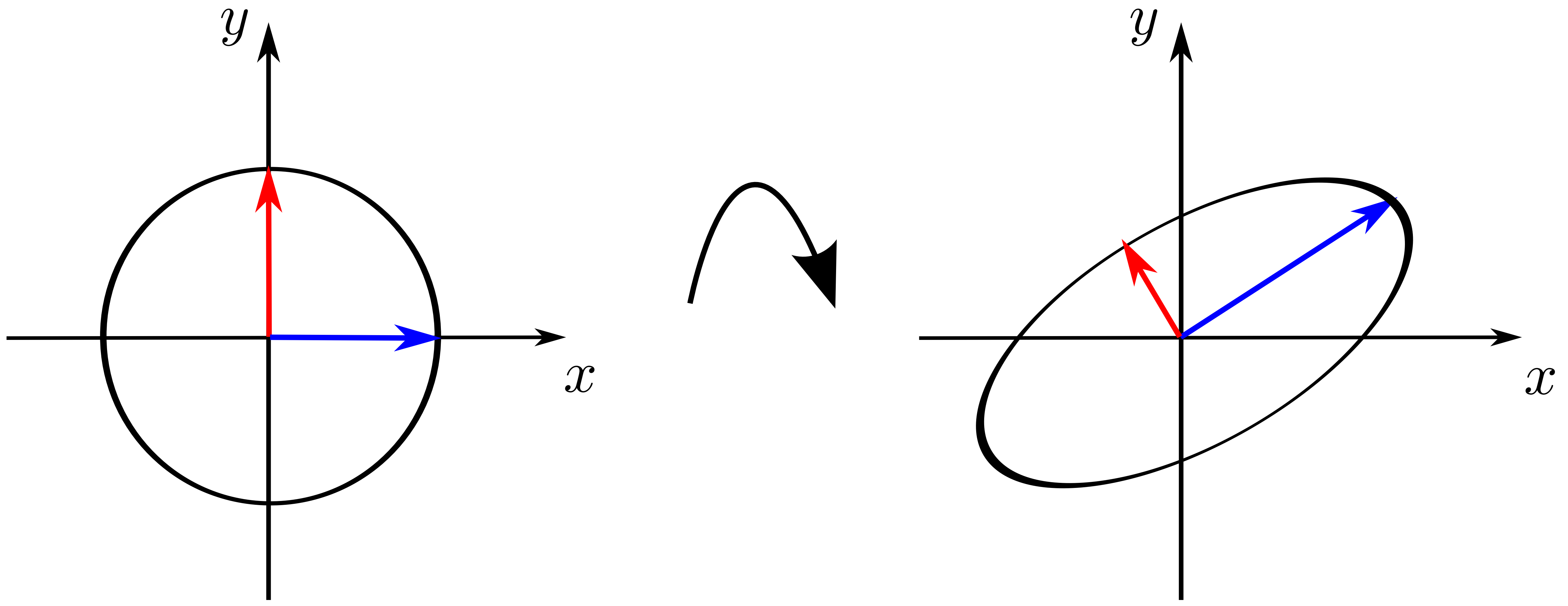


Figure 2: A rank 2-matrix $\boldsymbol{A}$ maps the 2-D unit sphere to a 2-D ellipse.

Normal Form of Matrices: SVD

Theorem

If $\boldsymbol{A}$ is a real $m \times n-$matrix, then there exist orthogonal matrices $\boldsymbol{U} \in \mathbb{R}^{m \times m}$ and $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ such that

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}},$$

where

$$\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{m \times n}$$

with $p = \min\{m, n\}$. The diagonal elements σ_i are the **singular values** *that fulfill*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0.$$

Take Home Messages

- SVD is a useful tool to solve a multitude of problems.
- We studied the effect of a matrix on unit vectors and the unit sphere.
- An arbitrary real matrix $\boldsymbol{A}$ can be decomposed by $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}}$.

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).

Medical Image Processing for Interventional Applications

Properties of the SVD

Online Course – Unit 3

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

Properties of the SVD: Rank and Norm

The singular value decomposition shows many extremely useful properties that are listed here without proof:

- ***rank*** of matrix $\boldsymbol{A}$:

$$\operatorname{rank}(\boldsymbol{A}) = \#\{\sigma_i > 0\},$$

- ***numerical*** ε***-rank*** of matrix $\boldsymbol{A}$:

$$\operatorname{rank}_\varepsilon(\boldsymbol{A}) = \#\{\sigma_i > \varepsilon\},$$

- the ***Frobenius norm*** of the matrix $\boldsymbol{A} = (a_{i,j})_{i,j}$ is given by

$$\|\boldsymbol{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2}.$$

Properties of the SVD: Eigenvectors

The singular value decomposition shows many extremely useful properties that are listed here without proof:

- decomposition into rank $1-$matrices:

$$\boldsymbol{A} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}, \quad r = \mathrm{rank}(\boldsymbol{A}),$$

- $\boldsymbol{A}\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{u}_i = \sigma_i \boldsymbol{v}_i$,
- the column vectors of $\boldsymbol{U}$ are the eigenvectors of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$:

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{u}_i = \sigma_i^2 \boldsymbol{u}_i,$$

- the column vectors of $\boldsymbol{V}$ are the eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$:

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_i = \sigma_i^2 \boldsymbol{v}_i.$$

Properties of the SVD

- The SVD yields orthonormal bases for the kernel (null-space) and the range of a matrix $\boldsymbol{A}$:
  - The ***kernel*** of $\boldsymbol{A}$ is spanned by the column vectors $\boldsymbol{v}_i$ of $\boldsymbol{V}$, where the corresponding singular values fulfill $\sigma_i = 0$.
  - The ***range*** of $\boldsymbol{A}$ is spanned by the column vectors $\boldsymbol{u}_i$ of $\boldsymbol{U}$, where σ_i are the corresponding non–zero singular values.
- For the 2-norm of matrix $\boldsymbol{A}$ we get:

$$\|\boldsymbol{A}\|_2^2 = \max_{\|\boldsymbol{x}\|_2=1} \boldsymbol{x}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\boldsymbol{x} = \sigma_1^2,$$

  and if $\boldsymbol{A}$ is regular, we even have:

$$\|\boldsymbol{A}^{-1}\|_2^2 = \frac{1}{\sigma_p^2}.$$

Example

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

where

$$\boldsymbol{U} = \begin{pmatrix} 0.1285 & 0.8375 & 0.5311 \\ -0.2396 & 0.5459 & -0.8028 \\ -0.9623 & -0.0241 & 0.2708 \end{pmatrix},$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 71.3967 & 0 & 0 \\ 0 & 21.7831 & 0 \\ 0 & 0 & 0.0006 \end{pmatrix},$$

$$\boldsymbol{V} = \begin{pmatrix} -0.2092 & 0.7082 & -0.6743 \\ -0.1941 & 0.6458 & 0.7384 \\ 0.9584 & 0.2854 & 0.0024 \end{pmatrix}.$$

Example

- Obviously, matrix $\boldsymbol{A}$ has a rank deficiency if we select $\varepsilon = 10^{-3}$.
- The kernel of $\boldsymbol{A}$ is given by:

$$\ker(\boldsymbol{A}) = \left\{ \lambda \cdot \begin{pmatrix} -0.6743 \\ 0.7384 \\ 0.0024 \end{pmatrix} ; \lambda \in \mathbb{R} \right\}.$$

- The range (or image) of $\boldsymbol{A}$ is:

$$\mathrm{im}(\boldsymbol{A}) = \left\{ \lambda \cdot \begin{pmatrix} 0.1285 \\ -0.2396 \\ -0.9623 \end{pmatrix} + \mu \cdot \begin{pmatrix} 0.8375 \\ 0.5459 \\ -0.0241 \end{pmatrix} ; \lambda, \mu \in \mathbb{R} \right\}.$$

Ill-conditioned Matrix

Definition

A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is called ***ill-conditioned*** if for a given linear system

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

minor changes in $\boldsymbol{b} \in \mathbb{R}^m$ cause major changes in $\boldsymbol{x} \in \mathbb{R}^n$.

Definition

The ***condition number*** of a regular matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ with respect to a matrix norm $\|.\|$ is defined by

$$\kappa(\boldsymbol{A}) = \left\|\boldsymbol{A}^{-1}\right\| \cdot \|\boldsymbol{A}\| .$$

If $\boldsymbol{A}$ is singular, $\kappa(\boldsymbol{A}) = +\infty$.

Ill-conditioned Matrix: Remarks

- A matrix with $\kappa(\boldsymbol{A})$ close to 1 is called ***well-conditioned***.
- A matrix with $\kappa(\boldsymbol{A})$ significantly greater than 1 is said to be ***ill-conditioned***.
- The condition number is a measure of the stability or sensitivity of a matrix.
- Using the 2-norm, the condition number of a regular matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ can be computed by SVD:

$$\kappa(\boldsymbol{A}) = \frac{\sigma_1}{\sigma_n},$$

  where σ_1 is the largest, and σ_n is the smallest singular value.
- The SVD allows for the exact computation of the condition number, but this is computationally expensive.

Ill-conditioned Matrix

Example

Consider the previous matrix

$$\boldsymbol{A} = \begin{pmatrix} 11 & 10 & 14 \\ 12 & 11 & -13 \\ 14 & 13 & -66 \end{pmatrix},$$

where we have $\det(\boldsymbol{A}) = 1$. The singular value decomposition of $\boldsymbol{A}$ results in the singular values:

$$\sigma_1 = 71.3967,\ \sigma_2 = 21.7831,\ \text{and}\ \sigma_3 = 0.0006.$$

Thus the condition number is $\kappa(\boldsymbol{A}) = 118994.5 \gg 1$.

Exercise: Show that a variation in $\boldsymbol{b}$ by 0.1% implies a change in $\boldsymbol{x}$ by 240%.

Take Home Messages

- We learned about important properties of the SVD, like
  - analytical and numerical rank definition,
  - Frobenius norm and 2-norm,
  - the relation between $\boldsymbol{U}$, $\boldsymbol{V}$ and the eigenvectors of $\boldsymbol{A}\boldsymbol{A}^{\mathsf{T}}$, $\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}$,
  - the relation between the kernel/range of $\boldsymbol{A}$ and the columns of $\boldsymbol{V}$, $\boldsymbol{U}$.
- For every matrix a condition number can be computed. Ill-conditioned matrices are numerically rather instable.

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`

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 $\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^{\mathrm{T}}$, then

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

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:

$$\sum_{i=1}^{4} \left(\boldsymbol{b}_i^T \boldsymbol{X} \boldsymbol{b}_i\right)^2 \rightarrow \min, \qquad \text{subject to} \qquad \|\boldsymbol{X}\|_F = 1,$$

$$\Leftrightarrow \qquad \boldsymbol{b}_i^T \boldsymbol{X} \boldsymbol{b}_i = 0, \qquad i = 1, \dots, 4, \qquad \|\boldsymbol{X}\|_F = 1.$$

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:

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

where the ***measurement matrix M*** is built from single elements of the sum.

Optimization Problem II

Example

Let us consider the i-th component:

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

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`

Medical Image Processing for Interventional Applications

SVD in Optimization - Part 2

Online Course – Unit 5

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

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Optimization Problem III

Another quite important optimization problem in image processing and pattern recognition is the following:

Problem: Given a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$.
Compute the matrix $\widehat{\boldsymbol{B}} \in \mathbb{R}^{n \times n}$ of rank $k < n$ that minimizes:

$$\widehat{\boldsymbol{B}} = \underset{\boldsymbol{B}}{\arg\min} \, \|\boldsymbol{A} - \boldsymbol{B}\|_2, \quad \text{subject to} \quad \operatorname{rank}(\boldsymbol{B}) = k.$$

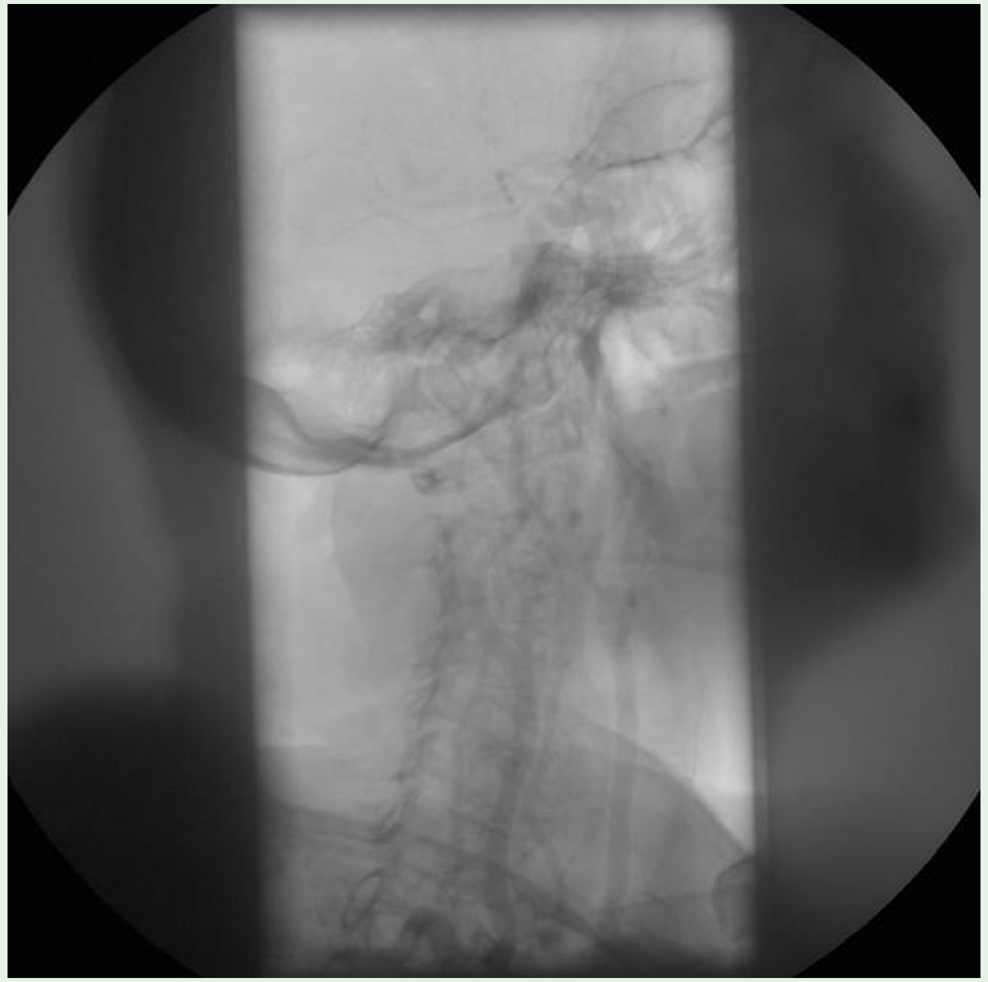
Solution: Using SVD, the solution can be computed easily by:

$$\widehat{\boldsymbol{B}} = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathsf{T}}.$$

Optimization Problem III

Example

The SVD can be used to compute the image matrix of rank 1 that approximates an image best w. r. t. $\|.\|_2$.
Figure 1 shows an image $\boldsymbol{I}$ and its rank 1 - approximation $\boldsymbol{I}' = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathsf{T}}$.



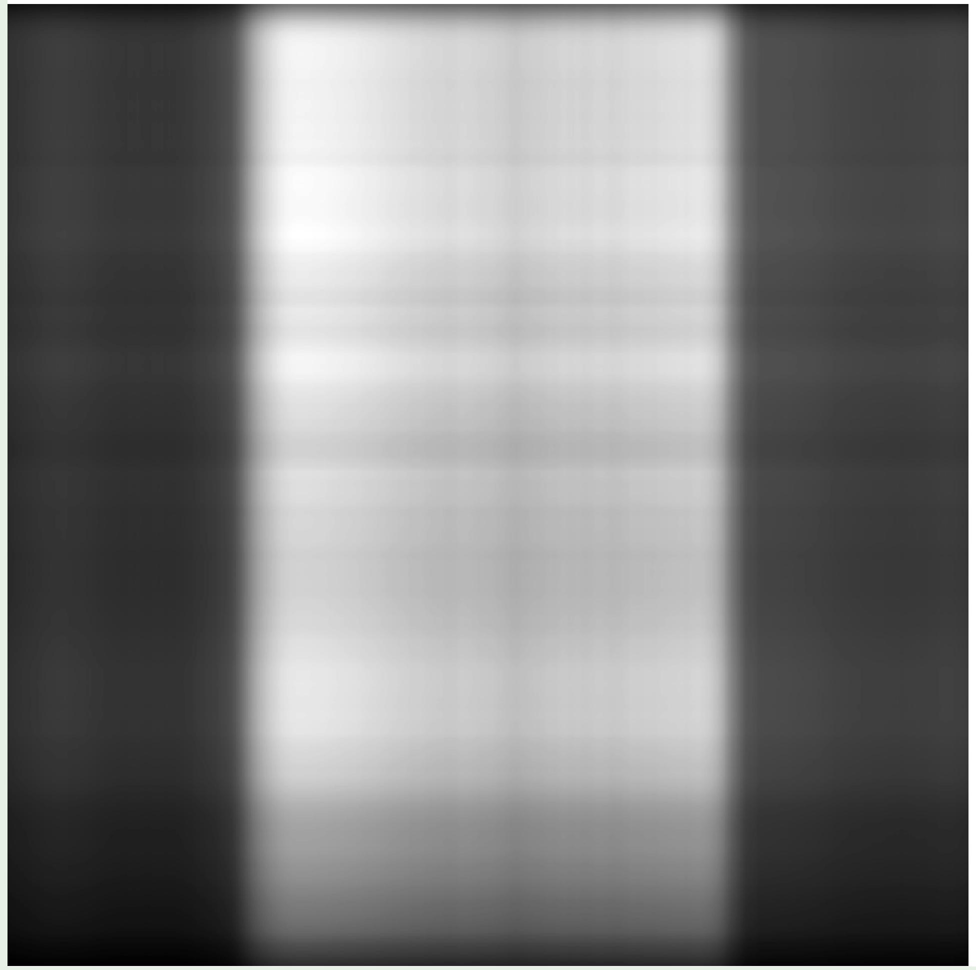


Figure 1: Original X-ray image (left) and its rank 1 - approximation (right)

Optimization Problem IV

Problem: The ***Moore–Penrose pseudoinverse*** is required to find the solution to the following optimization problem:

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2 \rightarrow \min .$$

Solution: The least squares solution of this optimization problem is given by

$$\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{b},$$

where we get $\boldsymbol{A}^{\dagger} \in \mathbb{R}^{n \times m}$ based on the SVD of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ by:

$$\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathsf{T}} = \boldsymbol{V}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{U}^{\mathsf{T}}.$$

Optimization Problem IV

Derivation: We start with the optimization problem:

$$\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 \rightarrow \min,$$

which can be solved analytically by derivation of this functional:

$$\begin{aligned} & \boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\right) = 0 \\ \Leftrightarrow \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} = 0 \\ \Leftrightarrow \quad & \boldsymbol{x} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}. \end{aligned}$$

Optimization Problem IV

The diagonal matrix $\boldsymbol{\Sigma}^\dagger$ in the SVD of the pseudo-inverse of $\boldsymbol{A}$ is given by:

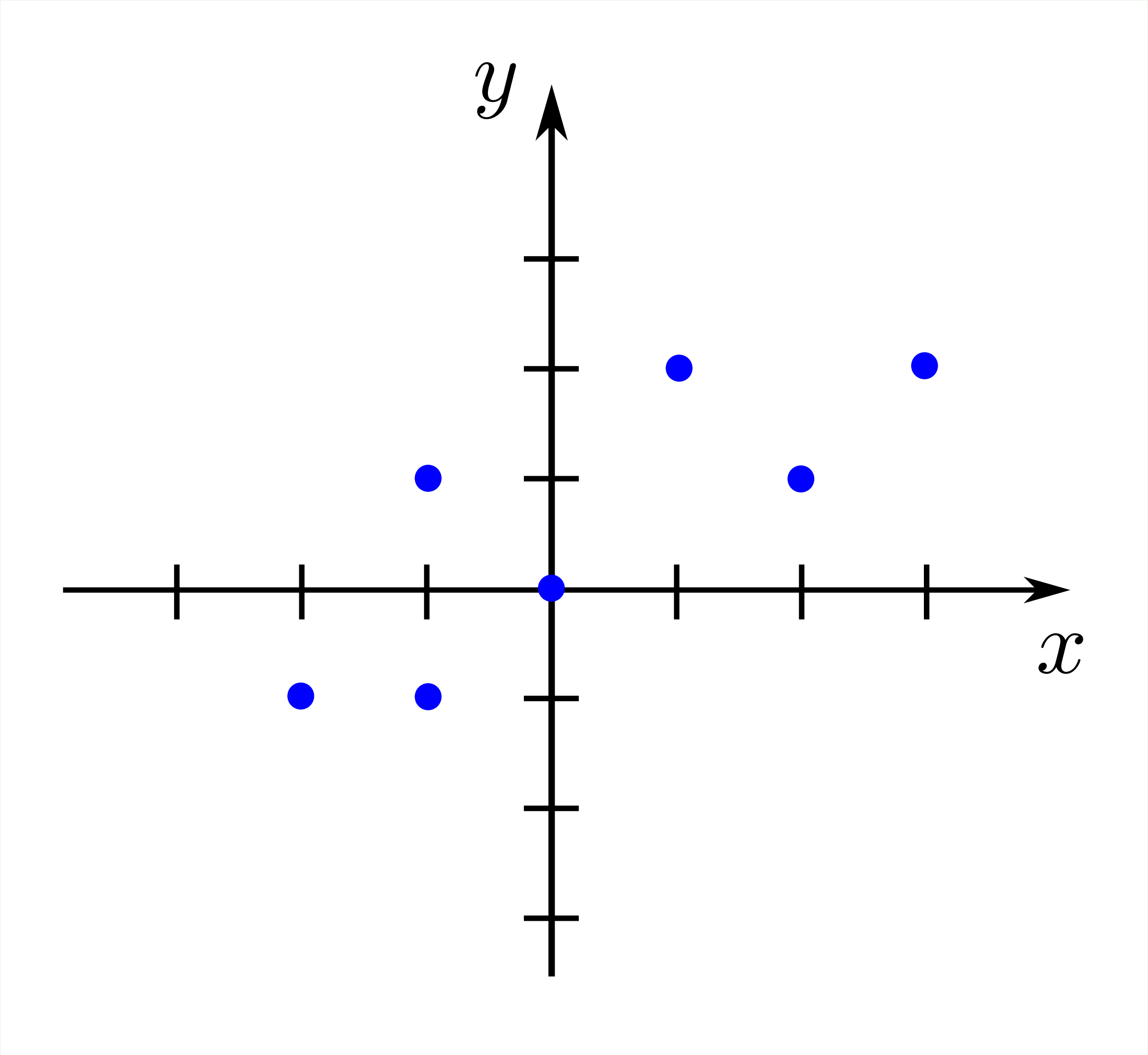
$$\boldsymbol{\Sigma}^\dagger = \begin{pmatrix} \frac{1}{\sigma_1} & & & & & 0 & \dots & 0 \\ & \ddots & & & & & & \\ & & \frac{1}{\sigma_r} & & & \vdots & & \vdots \\ & & & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{n \times m},$$

where $\sigma_r > 0$ is the smallest nonzero singular value of $\boldsymbol{A}$.

Optimization Problem IV

Example

Compute the regression line through the following 2-D points:



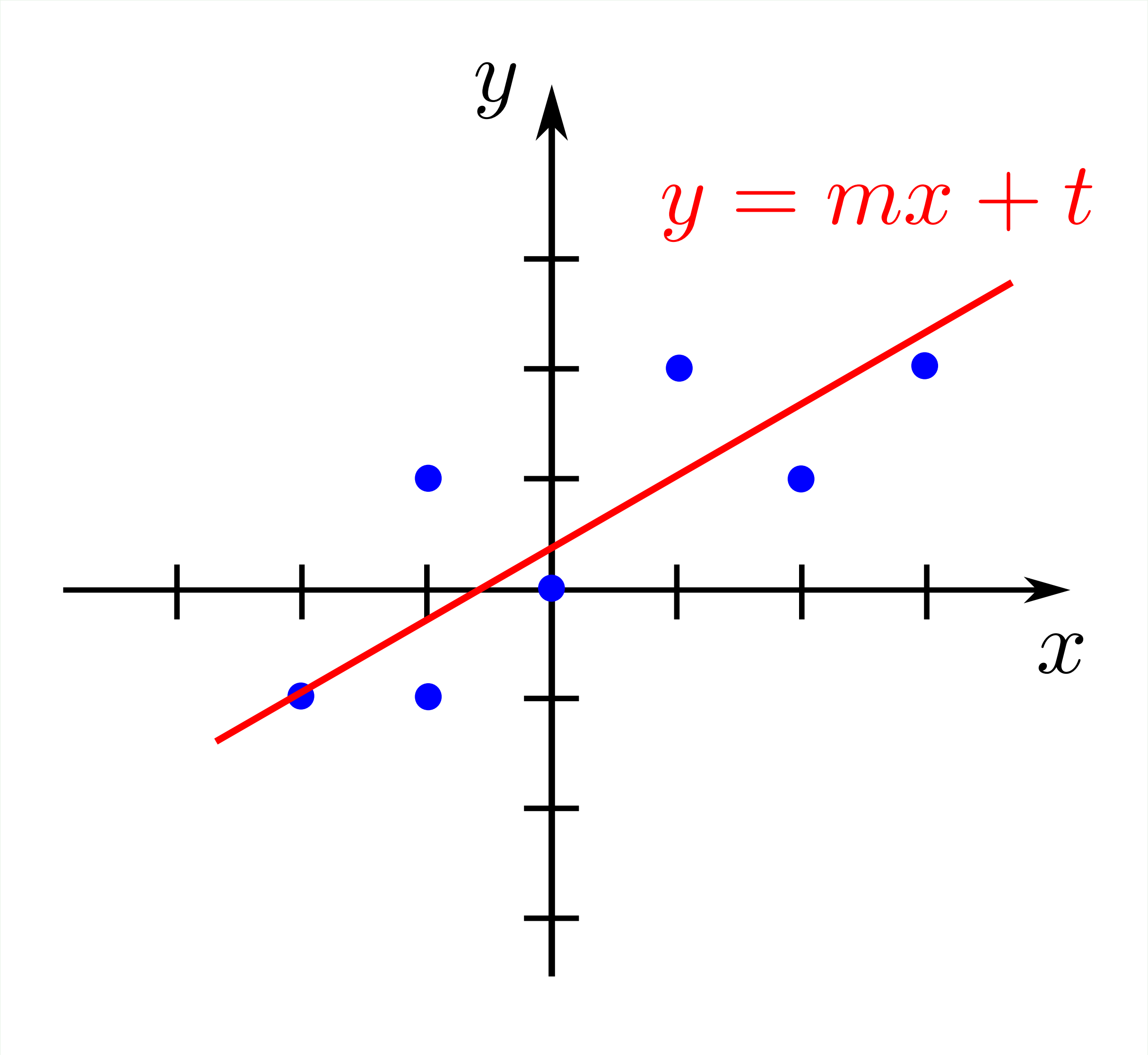


Figure 2: Regression line through a set of 2-D points

Optimization Problem IV

All points (x_i, y_i), $i = 1, \dots, 7$, have to fulfill the line equation:

$$y_i = mx_i + t, \quad \text{for} \quad i = 1, \dots, 7.$$

Thus we get the following system of linear equations:

$$\begin{pmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 1 \\ 0 & 1 \\ -1 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} m \\ t \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \\ 1 \\ -1 \\ -1 \end{pmatrix}.$$

Optimization Problem IV

The Moore-Penrose pseudo-inverse for this particular problem is:

$$\boldsymbol{A}^{\dagger} = \begin{pmatrix} 0.14 & 0.09 & 0.04 & -0.01 & -0.07 & -0.07 & -0.12 \\ 0.11 & 0.12 & 0.13 & 0.15 & 0.16 & 0.16 & 0.18 \end{pmatrix}.$$

Therefore, for the regression line we get the equation:

$$y = 0.56x + 0.41.$$

Remarks on SVD Computation

- SVD can be computed for every matrix.
- SVD is numerically robust.
- In most practical situations we have more rows than columns:

$$m \gg n.$$

- The time complexity to decompose $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is:

$$4m^2n + 8mn^2 + 9n^3.$$

- For us, the SVD is a black box. We do not consider algorithms to compute the SVD numerically.

Take Home Messages

- We have studied two additional applications (see also previous unit):
  - low-rank approximation,
  - fitting of a regression line.
- SVD is *the* tool for linear equations – it cannot fail (but in many special cases there may exist better solutions).
- SVD is provided by all standard libraries.

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`