Diagnostic Medical Image Processing

Singular Value Decomposition (SVD)

WS 2010/11

Joachim Hornegger

Lehrstuhl für Mustererkennung (Informatik 5)

Friedrich-Alexander-Universität Erlangen-Nürnberg

1 Singular Value Decomposition (SVD)
- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Singular Value Decomposition

$A = U \Sigma V^T$

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

Singular Value Decomposition

SVD is a perfect tool for the

- computation of singular values
- computation of null space
- computation of (pseudo-) inverse
- solution of overdetermined linear equations
- computation of condition numbers
- enforcing rank criterion (numerical rank)
- etc.

$A\vec{x} = \vec{0}$

$[\vec{a}_1, \vec{a}_2, \vec{a}_3]$

Diagnostic Medical Image Processing

1 Singular Value Decomposition (SVD)
- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

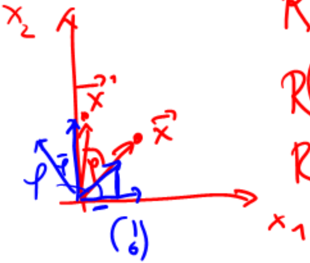
On the Geometry of Linear Mappings

From linear algebra, we know that a matrix $\boldsymbol{A}$ maps the unit vectors of the standard base to the corresponding column vector of matrix $\boldsymbol{A}$.

Example

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

Example:



$$R\vec{x} = \vec{x}'$$

$$R\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix}$$

$$R\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-\sin\varphi\\ \cos\varphi\end{pmatrix}$$

$$R = \begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

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 θ.

Solution:

The base vectors are mapped as follows:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \tag{2}$$

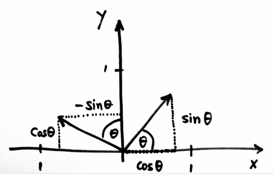


Figure 1: Rotation of 2-D unit vectors

and thus the 2-D rotation matrix is:

$$\boldsymbol{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3}$$

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 hyper-ellipsoid in the m-dimensional space.

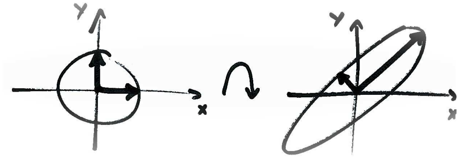


Figure 2: Rank 2 Matrix $\boldsymbol{A}$ maps 2-D unit sphere to the 2-D ellipse

1 Singular Value Decomposition (SVD)
- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

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*

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

where

$$\boldsymbol{\Sigma} \quad = \quad diag(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{m \times n} \tag{4}$$

with $p = \min(m, n)$*; the diagonal elements* σ_i *are the* **singular values** *that fulfill*

$$\sigma_1 \quad \geq \quad \sigma_2 \geq \cdots \geq \sigma_p \geq 0. \tag{5}$$

1 Singular Value Decomposition (SVD)

- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Properties of SVD: Rank, Norm and Eigenvectors

The Singular Value Decomposition shows many extremely useful properties that are listed here without proof:

- Rank of matrix $\boldsymbol{A}$: $\text{rank}(\boldsymbol{A}) = \#\{\sigma_i > 0\} = r$
- Numerical ϵ-rank of matrix $\boldsymbol{A}$: $\text{rank}_\epsilon(\boldsymbol{A}) = \#\{\sigma_i > \epsilon\}$
- $\boldsymbol{A} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathsf{T}}$
- The Frobenius norm of the matrix $\boldsymbol{A}$ is given by

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

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

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

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

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

Properties of SVD

- The SVD yields orthonormal bases for the kernel (null-space) and the range of a matrix $\boldsymbol{A}$:
  - The kernel of matrix $\boldsymbol{A}$ is spanned by the column vectors $\boldsymbol{v}_i$ of $\boldsymbol{V}$, where the corresponding singular values fulfill $\sigma_i = 0$.
  - The range of matrix $\boldsymbol{A}$ 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

$$\boldsymbol{A} = \begin{pmatrix} 11 & 10 & 14 \\ 12 & 11 & -13 \\ 14 & 13 & -66 \end{pmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}} \tag{6}$$

where

$$\boldsymbol{U} = \begin{pmatrix} 0.1285 & 0.8375 & 0.5311 \\ -0.2396 & 0.5459 & -0.8028 \\ -0.9623 & -0.0241 & 0.2708 \end{pmatrix} \tag{7}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 71.3967 & 0 & 0 \\ 0 & 21.7831 & 0 \\ 0 & 0 & 0.0006 \end{pmatrix} \tag{8}$$

$$\boldsymbol{V} = \begin{pmatrix} -0.2092 & 0.7082 & -0.6743 \\ -0.1941 & 0.6458 & 0.7384 \\ 0.9584 & 0.2854 & 0.0024 \end{pmatrix} \tag{9}$$

Example

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

$$\text{kernel}(\boldsymbol{A}) = \left\{ k \cdot \begin{pmatrix} -0.6743 \\ 0.7384 \\ 0.0024 \end{pmatrix} ; k \in \mathbb{R} \right\}$$

- The range of $\boldsymbol{A}$ is:

$$\text{range}(\boldsymbol{A}) = \left\{ k \cdot \begin{pmatrix} 0.1285 \\ -0.2396 \\ -0.9623 \end{pmatrix} + l \cdot \begin{pmatrix} 0.8375 \\ 0.5459 \\ -0.0241 \end{pmatrix} ; k, l \in \mathbb{R} \right\}$$

Ill-conditioned Matrix

Definition

A matrix $\boldsymbol{A}$ is called **ill-conditioned** if for a given linear system

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

minor changes in $\boldsymbol{b}$ cause major changes in $\boldsymbol{x}$

Definition

The **condition number** of a regular matrix $\boldsymbol{A}$ with respect to a matrix norm $||.||$ is defined by

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

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

Ill-conditioned Matrix

Remarks

- A matrix with a condition number close to 1 is called **well-conditioned**.
- A matrix with a condition number 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 the quadratic matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ can be computed by SVD:

$$\kappa(\boldsymbol{A}) = \frac{\sigma_1}{\sigma_n} \tag{10}$$

  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 it 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 obviously 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 $\kappa(\boldsymbol{A}) = 118994.5$.

Exercise problem:

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

Diagnostic Medical Image Processing

1 Singular Value Decomposition (SVD)

- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Optimization Problem I

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

We computed a matrix $\boldsymbol{A}$ out of sensor data like an image. By theory the matrix $\boldsymbol{A}$ must have the singular values $\sigma_1, \sigma_2, \ldots, \sigma_p$, where $p = \min(m, n)$. Of course, in practice $\boldsymbol{A}$ does not fulfill this constraint.

Problem: What is the matrix $\boldsymbol{A}'$ that is closest to $\boldsymbol{A}$ (according to the Frobenius norm) and has the required singular values?

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

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

Optimization Problem I

Example

The measurements lead to the following matrix:

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

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 $\boldsymbol{A}$ according to the Frobenius norm and fulfills the above requirements is:

$$\boldsymbol{A}' = \boldsymbol{U}\,\mathrm{diag}((71.3967 + 21.7831)/2, (71.3967 + 21.7831)/2, 0)\boldsymbol{V}^{\mathrm{T}}$$

Diagnostic Medical Image Processing

1 Singular Value Decomposition (SVD)

- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Optimization Problem II

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

$$\widehat{\boldsymbol{x}} = \operatorname{argmin}_{\boldsymbol{x}} \boldsymbol{x}^{\mathsf{T}} \boldsymbol{A}^{\mathsf{T}} \boldsymbol{A} \boldsymbol{x} \quad \text{subject to} \quad ||\boldsymbol{x}||_2 = 1. \tag{11}$$

or in the extreme:

$$\boldsymbol{A}\boldsymbol{x} = 0 \quad \text{subject to} \quad ||\boldsymbol{x}||_2 = 1. \tag{12}$$

Solution: The solution can be constructed using the rightmost column of $\boldsymbol{V}$ (check this!).

Optimization Problem II

Example

Estimate the matrix $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ such that for the following 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 optimization problem gets solved:

$$\sum_{i=1}^{4} \boldsymbol{b}^{\mathrm{T}}{}_i \boldsymbol{A} \boldsymbol{b}_i \to \min \quad \text{s.t.} \quad ||\boldsymbol{A}||_F = 1$$

Optimization Problem II

Example

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

$$\boldsymbol{M}\boldsymbol{a} = \boldsymbol{M}\begin{pmatrix} a_{1,1} \\ a_{1,2} \\ a_{2,1} \\ a_{2,2} \end{pmatrix} = \boldsymbol{0} \quad \text{s.t.} \quad ||\boldsymbol{a}||_2 = 1$$

where the **measurement matrix $\boldsymbol{M}$** is built from single elements of the sum. Let us consider the i-th component:

$$\boldsymbol{b}^{\mathrm{T}}{}_i \boldsymbol{A}\boldsymbol{b}_i = \boldsymbol{b}^{\mathrm{T}}{}_i \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \boldsymbol{b}_i = (b_{i,1}^2, b_{i,1}b_{i,2}, b_{i,1}b_{i,2}, b_{i,2}^2)\begin{pmatrix} a_{1,1} \\ a_{1,2} \\ a_{2,1} \\ a_{2,2} \end{pmatrix}$$

Optimization Problem II

Example

Using this result we get for the overall measurement matrix:

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

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

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

1 Singular Value Decomposition (SVD)

- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Optimization Problem III

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

$$\widehat{\boldsymbol{B}} = \text{argmin}_{\boldsymbol{B}} ||\boldsymbol{A} - \boldsymbol{B}||_2 \quad \text{s.t.} \quad \text{rank}(\boldsymbol{B}) = k \quad .$$

Solution: Using SVD, the solution is quite simple:

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

Optimization Problem III

Example

The SVD can be used to compute the image matrix of rank 1 that best approximates an image. Figure **??** shows an example of an image $\boldsymbol{I}$ and its rank-1-approximation $\boldsymbol{I}' = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}}$.

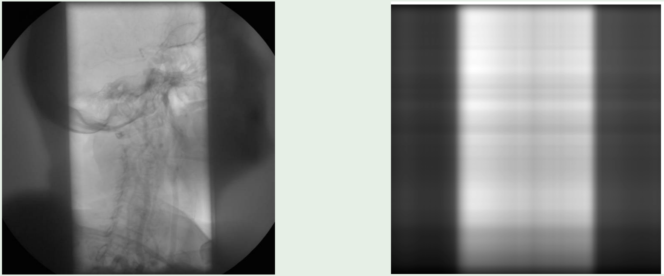


Figure 3: Original X-ray image and its rank-1-approximation

Diagnostic Medical Image Processing

1 Singular Value Decomposition (SVD)

- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Optimization Problem IV

Problem: The **Moore-Penrose pseudo-inverse** is required to find the solution to the following optimization problem:

$$||\boldsymbol{Ax} - \boldsymbol{b}||_2 \rightarrow 0 \tag{13}$$

Solution: The least square solution of this equation is given by

$$\boldsymbol{x} = \boldsymbol{A}^{\dagger}\boldsymbol{b} \tag{14}$$

where we get based on the SVD of $\boldsymbol{A}$:

$$\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathsf{T}} = \boldsymbol{V}\Sigma^{\dagger}\boldsymbol{U}^{\mathsf{T}}. \tag{15}$$

Optimization Problem IV

The diagonal matrix in the pseudo-inverse is defined via:

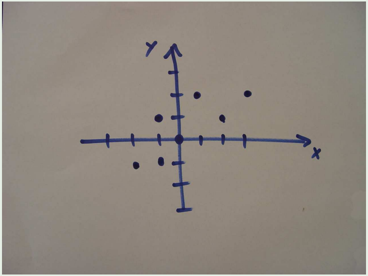
$$\boldsymbol{\Sigma}^{\dagger} = \begin{pmatrix} \frac{1}{\sigma_1} & & & & & 0 & \dots & 0 \\ & \ddots & & & & & & \\ & & \frac{1}{\sigma_r} & & & \vdots & & \vdots \\ & & & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{n \times m} \tag{16}$$

where $\sigma_r > \epsilon$ is the smallest nonzero singular value.

Optimization Problem IV

Example

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



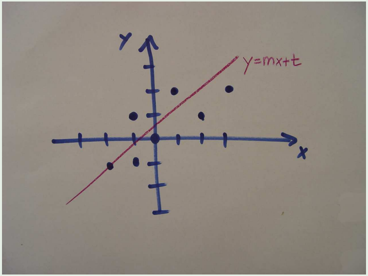


Figure 4: Regression line through 2-D points

Optimization Problem IV

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

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

Thus we get the 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:

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

The resulting line equation thus is:

$$y = 0.56x + 0.41$$

Remarks on SVD Computation

- For us SVD is a black box; we do not consider algorithms to compute the SVD numerically.
- SVD can be computed for any matrix
- SVD is numerically robust
- Time complexity to decompose $\boldsymbol{A} \in \mathbb{R}^{m \times n}$:

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

- In most practical situations we have more rows than columns, i.e. $m >> n$.

Diagnostic Medical Image Processing

1 Singular Value Decomposition (SVD)
- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Take Home Messages

- 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.
- SVD is always our first choice.
- SVD is most probably the right answer to any question in the oral exam. **Give it a try and check its limitations!**

Diagnostic Medical Image Processing

1 Singular Value Decomposition (SVD)
- General Remarks
- On the Geometry of Linear Mappings
- Normal Form of Matrices: SVD
- Properties of SVD
- Optimization Problem I
- Optimization Problem II
- Optimization Problem III
- Optimization Problem IV
- Take Home Messages
- Further Readings

Further Readings

- Read the original:
  Gene H. Golub, Charles F. Van Loan: Matrix Computations, Johns Hopkins Studies in Mathematical Sciences, 3rd edition, The Johns Hopkins University Press, Baltimore, 1996 (amazon the book right here).
- A very detailed and easy to follow introduction of the SVD can be found in Carlo Tomasi's class notes (download pdf, chapter 3) ... a **must-read**.
- The theory is described in an easy to read format (one of my favorite books!):
  Lloyd N. Trefethen, David Bau: Numerical Linear Algebra, Cambridge University Press, Cambridge, 1997 (amazon the book right here).

Further Readings

- For numerical computation of SVD see:
  William H. Press, Saul A. Teukolsky, William T. Vetterling: Numerical Recipes in C The Art of Scientific Computing, Cambridge University Press, Cambridge, 1993 (c.f. NR Web Page).