Singular Value Decomposition

Dr. D Bhanu Prakash
Course Page: dbhanuprakash233.github.io/LA

Assistant Professor,
Department of Mathematics and Statistics.
Contact: db_maths@vignan.ac.in.
dbhanuprakash233.github.io.

VIGNAN'S
FOUNDATION FOR SCIENCE, TECHNOLOGY & RESEARCH
(Deemed to be University) - Estd. u/s 3 of UGC Act 1956

Definition, Algorithm, Worked Example, and Applications

Outline

- Matrix Decomposition
- Singular Value Decomposition: Algorithm
- SVD: Worked example
- SVD: Applications to Image Compression

Outline

Matrix Decomposition: Definition & Importance

Definition (Matrix Decomposition / Factorization)

Given a matrix A, a decomposition writes

$$A = A_1 A_2 \cdots A_k$$

where factors A_i have a special structure (triangular, orthogonal, diagonal, etc.).

Matrix Decomposition: Definition & Importance

Definition (Matrix Decomposition / Factorization)

Given a matrix A, a decomposition writes

$$A = A_1 A_2 \cdots A_k$$

where factors A_i have a special structure (triangular, orthogonal, diagonal, etc.).

Why to decompose A?

- Turns a hard problem on A into easier problems on structured factors.

Matrix Decomposition: Definition & Importance

Definition (Matrix Decomposition / Factorization)

Given a matrix A, a decomposition writes

$$A = A_1 A_2 \cdots A_k$$

where factors A_i have a special structure (triangular, orthogonal, diagonal, etc.).

Why to decompose A?

- Turns a hard problem on A into easier problems on structured factors.
- Enables: solving linear systems, least squares, eigen-analysis, compression, stability analysis.

Matrix Decomposition: Definition & Importance

Definition (Matrix Decomposition / Factorization)

Given a matrix A, a decomposition writes

$$A = A_1 A_2 \cdots A_k$$

where factors A_i have a special structure (triangular, orthogonal, diagonal, etc.).

Why to decompose A?

- Turns a hard problem on A into easier problems on structured factors.
- Enables: solving linear systems, least squares, eigen-analysis, compression, stability analysis.
- Good decompositions give **numerical stability**, **interpretability**, and **computational efficiency**.

Common Decompositions

- **LU**: $A = LU$ (Gaussian elimination)
- **Cholesky**: $A = LL^\top$ for SPD matrices
- **QR**: $A = QR$ (least squares)
- **Eigendecomposition**: $A = Q\Lambda Q^{-1}$ (square matrices)

Common Decompositions

- **LU**: $A = LU$ (Gaussian elimination)
- **Cholesky**: $A = LL^\top$ for SPD matrices
- **QR**: $A = QR$ (least squares)
- **Eigendecomposition**: $A = Q\Lambda Q^{-1}$ (square matrices)

SVD (works for any $m \times n$)

$$A = U\Sigma V^\top$$

- Always exists (real/complex)
- Best low-rank approximation

Outline

Singular Value Decomposition

Singular Value Decomposition (Real case)

For any $A \in \mathbb{R}^{m\times n}$, there exist orthogonal matrices (i.e., $A^\top = A^{-1}$)

$$U \in \mathbb{R}^{m\times m}, \qquad V \in \mathbb{R}^{n\times n}$$

and a diagonal (rectangular) matrix

$$\Sigma \in \mathbb{R}^{m\times n}, \quad \Sigma = \begin{pmatrix} \mathrm{diag}(\sigma_1,\dots,\sigma_r) & 0 \\ 0 & 0 \end{pmatrix},$$

such that

$$A = U\Sigma V^\top, \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; r = \mathrm{rank}(A).$$

σ_i are **singular values**; columns of U are **left singular vectors**; columns of V are **right singular vectors**.

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m \times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n \times n}$.

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m \times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n \times n}$.
2. Solve eigenproblem: $A^\top A\, v_i = \lambda_i v_i$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m \times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n \times n}$.
2. Solve eigenproblem: $A^\top A\, v_i = \lambda_i v_i$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
3. Set singular values: $\sigma_i = \sqrt{\lambda_i}$, for $\lambda_i > 0$.

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m \times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n \times n}$.
2. Solve eigenproblem: $A^\top A\, v_i = \lambda_i v_i$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
3. Set singular values: $\sigma_i = \sqrt{\lambda_i}$, for $\lambda_i > 0$.
4. Form $V = [v_1 \;\; \cdots \;\; v_n]$ (orthogonal).

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m \times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n \times n}$.
2. Solve eigenproblem: $A^\top A\, v_i = \lambda_i v_i$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
3. Set singular values: $\sigma_i = \sqrt{\lambda_i}$, for $\lambda_i > 0$.
4. Form $V = [v_1 \ \cdots \ v_n]$ (orthogonal).
5. For each $\sigma_i > 0$: define $u_i = \dfrac{Av_i}{\sigma_i}$.

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m\times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n\times n}$.
2. Solve eigenproblem: $A^\top A\, v_i = \lambda_i v_i$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
3. Set singular values: $\sigma_i = \sqrt{\lambda_i}$, for $\lambda_i > 0$.
4. Form $V = [v_1 \;\; \cdots \;\; v_n]$ (orthogonal).
5. For each $\sigma_i > 0$: define $u_i = \dfrac{Av_i}{\sigma_i}$.
6. Complete $\{u_i\}$ to an orthonormal basis of $\mathbb{R}^m$ (add vectors spanning $\mathrm{null}(A^\top)$).

Step-by-Step SVD Construction (Mathematical Procedure)

Algorithm

Given $A \in \mathbb{R}^{m\times n}$:

1. Compute $A^\top A \in \mathbb{R}^{n\times n}$.
2. Solve eigenproblem: $A^\top A\, v_i = \lambda_i v_i$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
3. Set singular values: $\sigma_i = \sqrt{\lambda_i}$, for $\lambda_i > 0$.
4. Form $V = [v_1 \ \cdots \ v_n]$ (orthogonal).
5. For each $\sigma_i > 0$: define $u_i = \dfrac{Av_i}{\sigma_i}$.
6. Complete $\{u_i\}$ to an orthonormal basis of $\mathbb{R}^m$ (add vectors spanning $\mathrm{null}(A^\top)$).
7. Build Σ with σ_i on diagonal: $A = U\Sigma V^\top$.

Outline

Example Setup

Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

We will compute $A = U\Sigma V^\top$.

Example Setup

Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

We will compute $A = U\Sigma V^\top$.

Process

- Compute $A^\top A$ (a 2×2 symmetric matrix)
- Get eigenpairs (λ_i, v_i)
- $\sigma_i = \sqrt{\lambda_i}$, then $u_i = \dfrac{Av_i}{\sigma_i}$

Step 1: Compute $A^\top A$

Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

Step 1: Compute $A^\top A$

Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

$$A^\top = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \Rightarrow \quad A^\top A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Observation

$A^\top A$ is symmetric positive semidefinite $\Rightarrow$ real eigenvalues, orthonormal eigenvectors.

Step 2: Eigenvalues of $A^{\top}A$

Step 2: Eigenvalues of $A^\top A$

Solve

$$\det\left(\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix} - \lambda I\right) = \det\begin{pmatrix}2-\lambda & 1\\ 1 & 2-\lambda\end{pmatrix} = (2-\lambda)^2 - 1 = 0.$$

So

$$(2-\lambda)^2 = 1 \;\Rightarrow\; 2-\lambda = \pm 1 \;\Rightarrow\; \lambda_1 = 3,\; \lambda_2 = 1.$$

Singular values

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3}, \qquad \sigma_2 = \sqrt{\lambda_2} = 1.$$

Step 3: Right Singular Vectors (V)

Step 3: Right Singular Vectors (V)

For $\lambda_1 = 3$:

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - 3I\right) v = 0 \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} v = 0 \Rightarrow v_1 \propto \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Step 3: Right Singular Vectors (V)

For $\lambda_1 = 3$:

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - 3I\right) v = 0 \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} v = 0 \Rightarrow v_1 \propto \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Normalize:

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Step 3: Right Singular Vectors (V)

For $\lambda_1 = 3$:

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - 3I\right) v = 0 \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} v = 0 \Rightarrow v_1 \propto \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Normalize:

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For $\lambda_2 = 1$:

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - I\right) v = 0 \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v = 0 \Rightarrow v_2 \propto \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Step 3: Right Singular Vectors (V)

For $\lambda_1 = 3$:

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - 3I\right) v = 0 \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} v = 0 \Rightarrow v_1 \propto \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Normalize:

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For $\lambda_2 = 1$:

$$\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - I\right) v = 0 \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v = 0 \Rightarrow v_2 \propto \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Matrix V (columns are v_1, v_2)

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad V^\top V = I_2.$$

Step 4: Left Singular Vectors ($u_i = \frac{Av_i}{\sigma_i}$)

Step 4: Left Singular Vectors ($u_i = \frac{Av_i}{\sigma_i}$)

Compute Av_1:

$$Av_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Thus

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Step 4: Left Singular Vectors ($u_i = \frac{Av_i}{\sigma_i}$)

Compute Av_1:

$$Av_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Thus

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Compute Av_2:

$$Av_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Since $\sigma_2 = 1$,

$$u_2 = Av_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Step 5: Complete U and Build Σ

We need $U \in \mathbb{R}^{3\times 3}$ orthogonal. We already have u_1, u_2. Choose u_3 orthonormal to both (one valid choice):

$$u_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Assemble U

$$U = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix}, \qquad U^\top U = I_3.$$

Rectangular $\Sigma \in \mathbb{R}^{3\times 2}$

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Final SVD (Check)

Result

$$A = U\Sigma V^\top$$

with

$$V = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \quad \Sigma = \begin{pmatrix}\sqrt{3} & 0\\ 0 & 1\\ 0 & 0\end{pmatrix}, \quad U = \begin{pmatrix}\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}}\end{pmatrix}.$$

Quick sanity checks

$$\sigma_1 = \sqrt{3} \geq \sigma_2 = 1, \quad A^\top A = V\,\mathrm{diag}(3,1)V^\top, \quad AA^\top = U\,\mathrm{diag}(3,1,0)U^\top.$$

Outline

Applications of SVD: PCA, Denoising, Compression

Best rank-k approximation (Eckart–Young)

If $A = U\Sigma V^\top$ and $\Sigma_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$, then

$$A_k = U_k \Sigma_k V_k^\top$$

minimizes $\|A - A_k\|_2$ and $\|A - A_k\|_F$ among all rank-k matrices.

- **PCA**: principal components from SVD of centered data matrix
- **Denoising**: drop small singular values
- **Compression**: store U_k, Σ_k, V_k
- **Recommenders**: low-rank structure in user–item matrices

Least Squares & Pseudoinverse

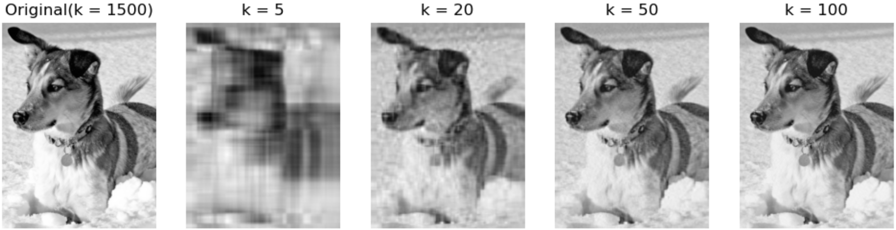
- Stable solutions to ill-conditioned problems (via truncation / regularization)
- Computing ranks, nullspaces, and condition numbers:

$$\kappa_2(A) = \frac{\sigma_1}{\sigma_r}.$$

SVD: Image Compression

Image Size: $1500 \times 1500 \times 3$

SVD: Grayscale Imaging



SVD: Color Imaging

Image Compression: Sizes & Storage Saved

Matrix size bookkeeping (grayscale)

Original image: $A \in \mathbb{R}^{m\times n}$ storage $\approx mn$ numbers.
Truncated SVD (rank k):

$$A \approx U_k \Sigma_k V_k^\top, \quad U_k \in \mathbb{R}^{m\times k},\ \Sigma_k \in \mathbb{R}^{k\times k},\ V_k \in \mathbb{R}^{n\times k}.$$

Storage $\approx mk + k + nk = k(m+n+1)$ numbers (since Σ_k is diagonal).

Space saved (when $k \ll \min(m,n)$)

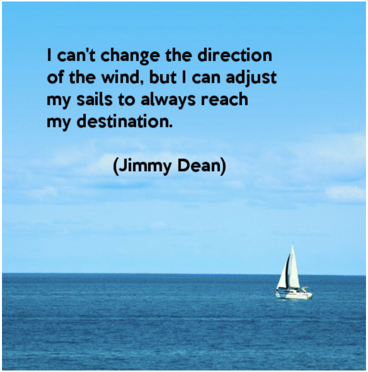
Compression ratio (approx.):

$$\mathrm{CR} \approx \frac{mn}{k(m+n+1)}.$$

?

Thank You!

Dr. D Bhanu Prakash
dbhanuprakash233.github.io
Mail: db_maths@vignan.ac.in



Applications I: Least Squares & Pseudoinverse

Least squares (possibly rank-deficient)

For $Ax \approx b$, the minimum-norm least squares solution is

$$x^\star = A^+ b, \qquad A^+ = V\Sigma^+ U^\top,$$

where Σ^+ replaces each nonzero σ_i by $1/\sigma_i$ and transposes dimensions.

- Stable solutions to ill-conditioned problems (via truncation / regularization)
- Computing ranks, nullspaces, and condition numbers:

$$\kappa_2(A) = \frac{\sigma_1}{\sigma_r}.$$

Applications II: PCA, Denoising, Compression, Recommendations

Best rank-k approximation (Eckart–Young)

If $A = U\Sigma V^\top$ and $\Sigma_k = \mathrm{diag}(\sigma_1, \dots, \sigma_k)$, then

$$A_k = U_k \Sigma_k V_k^\top$$

minimizes $\|A - A_k\|_2$ and $\|A - A_k\|_F$ among all rank-k matrices.

- **PCA**: principal components from SVD of centered data matrix
- **Denoising**: drop small singular values
- **Compression**: store U_k, Σ_k, V_k
- **Recommenders**: low-rank structure in user–item matrices

Rule of thumb

Energy captured by top-k:

$$\frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2}.$$

Geometric Meaning (Action on the Unit Sphere)

- $V^\top$ rotates/reflects in $\mathbb{R}^n$
- Σ stretches along coordinate axes
- U rotates/reflects in $\mathbb{R}^m$

Unit sphere to ellipsoid

For $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$, the set $\{Ax\}$ is an ellipsoid in $\mathbb{R}^m$ with semi-axis lengths $\sigma_1, \ldots, \sigma_r$.

Norm connections

$$\|A\|_2 = \sigma_1, \qquad \|A\|_F^2 = \sum_{i=1}^{r} \sigma_i^2.$$

Existence via $A^\top A$ and $AA^\top$

Core theorem (real case)

$A^\top A$ is symmetric positive semidefinite. Hence it has an orthonormal eigenbasis:

$$A^\top A = V\Lambda V^\top, \qquad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),\ \lambda_i \geq 0.$$

Define $\sigma_i = \sqrt{\lambda_i}$.

- Right singular vectors: eigenvectors of $A^\top A$
- Left singular vectors: eigenvectors of $AA^\top$

Constructing singular vectors

If v_i is an eigenvector of $A^\top A$ with $\lambda_i > 0$, then

$$u_i = \frac{Av_i}{\sigma_i}$$

is a unit left singular vector.

Computational Algorithm (Numerical Linear Algebra View)

Stable SVD (high-level steps)

Most software computes SVD without explicitly forming $A^\top A$ (for stability):

1. **Bidiagonalization**: $A = Q_1 B Q_2^\top$ where Q_1, Q_2 orthogonal and B bidiagonal.
2. **Diagonalization of B**: compute $B = \tilde{U}\Sigma\tilde{V}^\top$ using QR iterations on bidiagonal form.
3. Combine: $U = Q_1\tilde{U}$, $V = Q_2\tilde{V}$ so $A = U\Sigma V^\top$.

Why avoid $A^\top A$ numerically?

Condition numbers square: $\kappa(A^\top A) = \kappa(A)^2$ (loss of accuracy for ill-conditioned matrices).