Matrix Decompositions (SVD, QR, LU)

Table of Contents

Why Decompose?

Decompositions express a matrix as a product of simpler, structured matrices. This reveals geometry, improves numerical stability, and powers algorithms in QC/ML.

- Compression and denoising (SVD)
- Least-squares and orthogonalization (QR)
- Fast linear solves and reuse across right-hand sides (LU)

Singular Value Decomposition (SVD)

For any matrix $A \in \mathbb{R}^{m\times n}$, there exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$, and diagonal $\Sigma \in \mathbb{R}^{m\times n}$ with nonnegative entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ such that:

$$A = U\,\Sigma\,V^T.$$

- Columns of U are left singular vectors; columns of V are right singular vectors.
- Singular values σ_i are square-roots of eigenvalues of A^TA (or AA^T).
- Best rank-k approximation (Eckart–Young):

  $$A_k = U_{:,1:k}\,\Sigma_{1:k,1:k}\,V^T_{:,1:k}$$

  minimizes $\|A - A_k\|_F$.

Geometric view: V rotates coordinates, Σ scales along orthogonal axes, U rotates to final position.

Animation: Animation: In `interactive-visualizations.py`, the "Matrix Transformations" demo visualizes orthogonal axes and anisotropic scaling.

QR Decomposition

For $A \in \mathbb{R}^{m\times n}$ with full column rank, QR decomposes as:

$$A = QR,$$

where $Q \in \mathbb{R}^{m\times n}$ has orthonormal columns and $R \in \mathbb{R}^{n\times n}$ is upper triangular with positive diagonal.

- Gram–Schmidt yields Q and R.
- Solving least squares: $\min_x \|Ax - b\|_2$ becomes $Rx = Q^T b$.

LU Decomposition

For many square matrices, we can factor:

$$A = LU,$$

with L unit lower triangular and U upper triangular (possibly with row permutations $PA = LU$).

- Efficient for solving $Ax = b$ with repeated right-hand sides.
- Foundation for many direct solvers in numerical linear algebra.

QC / ML Parallels

- QC: Density matrices ρ (PSD, trace 1) admit eigen/SVD-like decompositions revealing mixed-state probabilities and purifications. Partial traces relate to block-structured SVD.
- ML: SVD underpins compression, noise reduction, and recommendation systems (latent factor models). QR appears in stable least-squares; LU in fast batched solves.

Worked Examples

Example 1: SVD of a 2×2 Matrix

Let $A = \begin{pmatrix} 3 & 0 \\ 4 & 0 \end{pmatrix}$. Then $A^T A = \begin{pmatrix} 25 & 0 \\ 0 & 0 \end{pmatrix}$ with eigenvalues $25, 0$. Thus $\sigma_1 = 5, \sigma_2 = 0$. One valid SVD is

$$U = \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix}, \ \Sigma = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}, \ V = I.$$

Example 2: QR via Gram–Schmidt

For $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, apply Gram–Schmidt to obtain Q with orthonormal columns and compute $R = Q^T A$.

Example 3: LU with Partial Pivoting

Factor $A = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}$. With permutation $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, we have $PA = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} = LU$ with $L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$.

Practice Problems

1. Show that singular values are nonnegative and equal to the square-roots of eigenvalues of $A^T A$.
2. Compute the best rank-1 approximation of $\begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.
3. Use QR to solve the least squares problem for $\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} x \approx \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$.
4. Perform LU factorization with partial pivoting for $\begin{pmatrix} 2 & 4 & 8 \\ 1 & 3 & 2 \\ 3 & 10 & 4 \end{pmatrix}$.
5. QC/ML: Explain how SVD underlies PCA and how eigen-decomposition of covariance relates to SVD of the data matrix.

Solutions

Solution to Problem 1

For any A, $A^T A$ is symmetric PSD. If $A^T A v = \lambda v$ with $\lambda \geq 0$, define $\sigma = \sqrt{\lambda} \geq 0$. Then singular values are $\sigma_i = \sqrt{\lambda_i(A^T A)} \geq 0$.

Solution to Problem 2

Let $B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. The best rank-1 approximation is $\sigma_1 u_1 v_1^T$ using the top singular triplet. Here $\sigma_1 = 2$, $u_1 = (1, 0, 0)^T$, $v_1 = (1, 0)^T$, so $B_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Solution to Problem 3

QR via Gram–Schmidt yields Q with orthonormal columns and $R = Q^T A$. Solving $Rx = Q^T b$ gives the least-squares solution $x = (1, 1)^T$ for the given data.

Solution to Problem 4

Apply Gaussian elimination with partial pivoting to produce P, then read off L (unit lower) and U (upper). One valid factorization is obtainable with standard steps; determinants of leading principal minors are nonzero so factorization exists with a suitable permutation.

Solution to Problem 5

For centered data matrix X (rows as samples), $\Sigma = \frac{1}{m} X^T X$. If $X = U \Sigma_X V^T$ is the SVD, then $X^T X = V \Sigma_X^2 V^T$; columns of V (right singular vectors) are the eigenvectors of Σ (principal components) and squared singular values are the eigenvalues (variances along components).