Matrix decompositions

Zdeněk Dvořák

May 19, 2015

Lemma 1 (Schur decomposition)**.** *If A is a symmetric real matrix, then there exists an orthogonal matrix Q and a diagonal matrix D such that $A = QDQ^T$. The diagonal entries of D are the eigenvalues of A.*

Lemma 2 (Cholesky decomposition)**.** *If A is a positive definite $n \times n$ matrix, then there exists a unique lower-triangular matrix L with positive entries on the diagonal such that $A = LL^T$.*

1 LU(P) decomposition

Definition 1. *Let A be an $n \times m$ matrix. Then $A = LU$ is an <u>LU decomposition</u> of A if L is a lower-triangular $n \times n$ matrix with ones on the diagonal, and U is an upper-triangular $n \times m$ matrix.*

Similar to Cholesky decomposition, but does not require positive semidefiniteness of A. LU decomposition does not always exist. But:

Lemma 3. *For every square matrix A, there exists a permutation matrix P such that PA has an LU decomposition.*

Proof. Reorder the rows of A so that Gaussian elimination algorithm does not need to exchange rows—the reordering is described by the permutation matrix P. Run Gaussian elimination for PA, only allowing addition of a multiple of a row to some row with higher index; the resulting matrix is U. The matrix L is obtained by performing the inverse operations on an identity matrix in the reverse order. □

Example 1. *Find an LUP decomposition of* $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \end{pmatrix}$.

In the Gaussian elimination algorithm, we select the first pivot in the 1st row, the second pivot in the 3rd row, and the third pivot in the 2nd row (and the remaining rows are zero, afterwards). Hence, we set $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, *so that* $PA = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 2 & 2 & 0 & 0 \\ 2 & 2 & 1 & 1 \end{pmatrix}$ *and the Gaussian elimination no longer needs to exchange rows.*

In the Gaussian elimination, we in order add

$\alpha\times$	*i-th row to*	*j-th row*
-2	1	3
-2	1	4
$-1/2$	3	4,

ending up with $U = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. *This corresponds to* $L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 2 & 0 & 1/2 & 1 \end{pmatrix}$. *Hence,*

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 2 & 0 & 1/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Applications:

- To solve system $Ax = b$, equivalently we can solve $PAx = LUx = Pb$ by forward and backward substitution (efficient when solving repeatedly with different right-hand sides).
- Computing inverse, determinant (slower than direct methods).

2 QR decomposition

Definition 2. *Let A be an $n \times m$ matrix. Then $A = QR$ is a QR decomposition of A if Q is an orthogonal $n \times n$ matrix and R is an upper-triangular $n \times m$ matrix.*

Lemma 4. *Every matrix has a QR decomposition.*

Proof. Run the Gram-Schmidt ortonormalization process for the columns v_1, $\dots$, v_m of A to obtain an orthonormal basis u_1, $\dots$, u_k of $\mathrm{span}(v_1, \dots, v_m)$, such that for $i = 1, \dots, k$, we have $v_i \in \mathrm{span}(u_1, \dots, u_i)$. Let u_{k+1}, $\dots$, u_n be arbitrary vectors extending u_1, $\dots$, u_k to an orthonormal basis, and let $Q = (u_1| \dots |u_n)$. Then $R = Q^{-1}A = Q^T A$ is upper-triangular. $\square$

Example 2. *Compute a QR decomposition of* $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \end{pmatrix}$.

Gram-Schmidt ortonormalization process for the columns of A gives

$$\frac{1}{3}(1, 2, 0, 2)^T, (0, 0, 1, 0)^T, 1/3(2, -2, 0, 1)^T.$$

This can be extended to an orthonormal basis by adding the vector $1/3(2, 1, 0, -2)$, *hence we can set* $Q = \begin{pmatrix} 1/3 & 0 & 2/3 & 2/3 \\ 2/3 & 0 & -2/3 & 1/3 \\ 0 & 1 & 0 & 0 \\ 2/3 & 0 & 1/3 & -2/3 \end{pmatrix}$.

Hence, $R = Q^T A = \begin{pmatrix} 3 & 3 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Remark: there exist better (numerically more stable) ways of computing QR decomposition.

Definition 3. *Let $A = QR$ be a QR decomposition of an $n \times m$ matrix A of rank k, where last $n - k$ rows of R are equal to 0. Let R' be the $k \times m$ matrix consisting of the first k rows of R, and let $Q = (Q'|Q'')$, where Q' has k columns. Then $A = Q'R'$ is a reduced QR decomposition of A.*

Applications:

- To solve system $Ax = b$, solve $Rx = Q^T b$ by substitution (numerically more stable than Gaussian elimination, but slower).
- If A is an $n \times m$ matrix with $m \leq n$, the rank of A is m, and $A = Q'R'$ is a reduced QR decomposition of A, then the solution to $R'x = (Q')^T b$ is the least-squares approximate solution to $Ax = b$.

- QR algorithm to compute eigenvalues (basic idea): Let $A_0 = A$ and for $i \geq 0$, let $A_i = Q_i R_i$ be a QR decomposition of A_i and let $A_{i+1} = R_i Q_i$. Note that $A_{i+1} = R_i Q_i = Q_i^{-1} Q_i R_i Q_i = Q_i^{-1} A_i Q_i$, and thus A_0, A_1, ... have the same eigenvalues. The sequence typically converges to an upper-triangular matrix, whose diagonal entries give the eigenvalues.

3 Singular value (SVD) decomposition

Definition 4. *Let A be an $n \times m$ matrix. Then $A = Q_1 D Q_2^T$ is an* <u>*SVD decomposition*</u> *of A if Q_1 is an orthogonal $n \times n$ matrix, D is a diagonal $n \times m$ matrix with non-negative entries on the diagonal, and Q_2 is an orthogonal $m \times m$ matrix. Let $r = rank(A)$. The non-zero entries $d_1, \ldots, d_r$ of the diagonal matrix D are positive, and we call them the* <u>*singular values*</u> *of A.*

Lemma 5. *If $d_1, \ldots, d_r$ are the singular values of A, then $d_1^2, \ldots, d_r^2$ are the non-zero eigenvalues of $A^T A$.*

Proof. We have $A^T A = (Q_2 D^T Q_1^T) Q_1 D Q_2^T = Q_2 (D^T D) Q_2^T$. Note that $D^T D$ is diagonal with $d_1^2, \ldots, d_r^2$ on the diagonal and that $A^T A$ and $D^T D$ have the same eigenvalues. □

To construct an SVD decomposition, first find a Schur decomposition $Q_2 D' Q_2^T$ of $A^T A$, thus determining Q_2 and the singular values $d_1, \ldots, d_r$ as the roots of the elements of the diagonal of D'; this also determines D and $r = \mathrm{rank}(A)$. Let Q_2' consist of the first r columns of Q_2, let D'' be the $r \times r$ diagonal matrix with $d_1^{-1}, \ldots, d_r^{-1}$ on the diagonal, and let $Q_1' = A Q_2' D''$. Extend the $n \times r$ matrix Q_1' to an orthonormal matrix Q_1 arbitrarily.

Example 3. *Compute an SVD decomposition of* $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \end{pmatrix}$.

We have $A^T A = \begin{pmatrix} 9 & 9 & 3 & 3 \\ 9 & 10 & 4 & 4 \\ 3 & 4 & 3 & 3 \\ 3 & 4 & 3 & 3 \end{pmatrix}$ *with Schur decomposition $Q_2 L Q_2^T$ for*

$$Q_2 = \begin{pmatrix} -0.615 & 0.473 & -0.631 & 0.000 \\ -0.673 & 0.101 & 0.732 & 0.000 \\ -0.290 & -0.619 & -0.181 & -0.707 \\ -0.290 & -0.619 & -0.181 & 0.707 \end{pmatrix} \text{ and } L = \begin{pmatrix} 21.670 & 0 & 0 & 0 \\ 0 & 3.059 & 0 & 0 \\ 0 & 0 & 0.272 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the singular values are $\sqrt{21.670} \approx 4.655$, $\sqrt{3.059} \approx 1.749$, *and* $\sqrt{0.272} \approx 0.522$.

We get $Q_1' = AQ_2'D'' = A \begin{pmatrix} -0.615 & 0.473 & -0.631 \\ -0.673 & 0.101 & 0.732 \\ -0.290 & -0.619 & -0.181 \\ -0.290 & -0.619 & -0.181 \end{pmatrix} \begin{pmatrix} 0.215 & 0 & 0 \\ 0 & 0.572 & 0 \\ 0 & 0 & 1.916 \end{pmatrix} = \begin{pmatrix} -0.402 & -0.380 & -0.500 \\ -0.554 & 0.657 & 0.387 \\ -0.269 & -0.650 & 0.709 \\ -0.679 & -0.051 & -0.307 \end{pmatrix}$. *Extending* Q_1' *to an orthonormal matrix, we get* $Q_1 = \begin{pmatrix} -0.402 & -0.380 & -0.500 & 0.666 \\ -0.554 & 0.657 & 0.387 & 0.334 \\ -0.269 & -0.650 & 0.709 & -0.050 \\ -0.679 & -0.051 & -0.307 & -0.665 \end{pmatrix}$.

Applications:

- Decomposes $f(x) = Ax$ to isometries and scaling (singular values describe the deformation).
- For regular matrix A, the ratio d_1/d_n of singular values estimates loss of precision for matrix computations (inverse, solution of systems of equations, ...).
- Signal processing, compression (replacing small singular values by 0 does not change the matrix too much).
- The Frobenius norm of a matrix A is $\|A\| = \sqrt{\sum_{i,j} A_{i,j}^2} = \sqrt{\mathrm{Trace}(A^T A)} = \sqrt{\sum_i d_i^2}$, where $d_1, \ldots, d_r$ are the singular values of A. Hence the smallest singular value of A is the distance in Frobenius norm from A to a singular matrix (and actually, there exists no closer singular matrix).