Singular Value Decomposition

Notes on Linear Algebra

Chia-Ping Chen

Department of Computer Science and Engineering
National Sun Yat-Sen University
Kaohsiung, Taiwan ROC

Introduction

- The singular value decomposition, SVD, is just as amazing as the LU and QR decompositions.
- It is closely related to the diagonal form $A = Q\Lambda Q^T$ of a symmetric matrix. What happens if the matrix is not symmetric?
- It turns out that we can factorize A by $Q_1 \Sigma Q_2^T$, where Q_1, Q_2 are orthogonal and Σ is nonnegative and diagonal-like. The diagonal entries of Σ are called the singular values.

SVD Theorem

- Any $m \times n$ real matrix A can be factored into

$$A = Q_1 \Sigma Q_2^T = \text{(orthogonal)(diagonal)(orthogonal).}$$

- The matrices are constructed as follows: The columns of Q_1 $(m \times m)$ are the eigenvectors of AA^T, and the columns of Q_2 $(n \times n)$ are the eigenvectors of A^TA. The r singular values on the diagonal of Σ $(m \times n)$ are the square roots of the nonzero eigenvalues of both AA^T and A^TA.

Proof of SVD Theorem

The matrix $A^T A$ is real symmetric so it has a complete set of orthonormal eigenvectors: $A^T A x_j = \lambda_j x_j$, and

$$x_i^T A^T A x_j = \lambda_j x_i^T x_j = \lambda_j \delta_{ij}.$$

For positive λ_j's (say $j = 1, \ldots, r$), we define $\sigma_j = \sqrt{\lambda_j}$ and $q_j = \frac{Ax_j}{\sigma_j}$. Then $q_i^T q_j = \delta_{ij}$. Extend the q_i's to a basis for R^m. Put x's in Q_2 and q's in Q_1, then

$$(Q_1^T A Q_2)_{ij} = q_i^T A x_j = \begin{cases} 0 & \text{if } j > r, \\ \sigma_j q_i^T q_j = \sigma_j \delta_{ij} & \text{if } j \leq r. \end{cases}$$

That is, $Q_1^T A Q_2 = \Sigma$. So $A = Q_1 \Sigma Q_2^T$.

Remarks

- For positive definite matrices, SVD is identical to $Q\Lambda Q^T$. For indefinite matrices, any negative eigenvalues in Λ become positive in Σ.
- The columns of Q_1, Q_2 give orthonormal bases for the fundamental subspaces of A. (Recall that the nullspace of A^TA is the same as A).
- $AQ_2 = Q_1\Sigma$, meaning that A multiplied by a column of Q_2 produces a multiple of column of Q_1.
- $AA^T = Q_1\Sigma\Sigma^TQ_1^T$ and $A^TA = Q_2\Sigma^T\Sigma Q_2^T$, which mean that Q_1 must be the eigenvector matrix of AA^T and Q_2 must be the eigenvector matrix of A^TA.

Applications of SVD

- Through SVD, we can expand a matrix to be a sum of rank-one matrices

$$A = Q_1 \Sigma Q_2^T = u_1 \sigma_1 v_1^T + \cdots + u_r \sigma_r v_r^T .$$

- Suppose we have a 1000×1000 matrix, for a total of 10^6 entries. Suppose we use the above expansion and keep only the 50 most most significant terms. This would require $50(1 + 1000 + 1000)$ numbers, a save of space of almost 90%.
- This is used in image processing and information retrieval (e.g. Google).

SVD for Image

A picture is a matrix of gray levels. This matrix can be approximated by a small number of terms in SVD.

Pseudoinverse

- Suppose $A = Q_1 \Sigma Q_2^T$ is the SVD of an $m \times n$ matrix A. The pseudoinverse of A is defined by

$$A^+ = Q_2 \Sigma^+ Q_1^T,$$

where Σ^+ is $n \times m$ with diagonals $\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}$.

- The pseudoinverse of A^+ is A, or $A^{++} = A$.
- The minimum-length least-square solution to $Ax = b$ is $x^+ = A^+ b$. This is a generalization of the least-square problem when the columns of A are not required to be independent.

Proof of Minimum Length

Multiplication by Q_1^T leaves the length unchanged, so

$$|Ax - b| = |Q_1 \Sigma Q_2^T x - b| = |\Sigma Q_2^T x - Q_1^T b| = |\Sigma y - Q_1^T b|,$$

where $y = Q_2^T x = Q_2^{-1} x$. Since Σ is a diagonal matrix, we know the minimum-length least-square solution is $y^+ = \Sigma^+ Q_1^T b$. Since $|y| = |x|$, the minimum-length least-square solution for x is

$$x^+ = Q_2 y^+ = Q_2 \Sigma Q_1^T b = A^+ b.$$