Singular Value Decomposition, briefly

In a nutshell, the SVD is very simple. There is a matrix A. In the case of the indices computation it is a table of object-feature measure data. The objects corresponds to rows while the features corresponds to columns. Also, the matrix A is a linear operator. It maps a weights vector $\mathbf{w}$ in the weights space $\mathbb{R}^m$ to an indices vector $\mathbf{q}$ in the indices space $\mathbb{R}^n$. Here m is the number of objects and n is the number of features. A linear operator A can be represented as the product of three linear operators, $A = U\Lambda V^T$. The matrix U and V are orthogonal and the matrix Λ is diagonal. So, an arbitrary linear operator A could be represented as the product of a rotation, scaling and rotation linear operators. This quality of the SVD will be used in the indices computation algorithm. Below we will discuss the SVD in detail.

An arbitrary matrix $A = \{a_{ij}\}$ can be described as

$$a_{ij} = \sum_{k=1}^{r} u_{ik}\lambda_k v_{kj} + c_{ij}, \qquad (1)$$

where $i = 1, ..., m$ и $j = 1, ..., n$. Values of u_{kj}, λ_k and v_{jk} for given k one can obtain from the minimum of ε_n^2, where

$$\varepsilon_n^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}^2, \qquad (2)$$

with conditions of the normalization

$$\sum_{j=1}^{n} u_{kj}^2 = \sum_{i=1}^{m} v_{ik}^2 = 1 \qquad (3)$$

and the order $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_r \geq ... \geq 0$.

Rewrite (1), (2) and (3) in matrix notations:

$$\begin{aligned} &A = U\Lambda V^T + C, \\ &\varepsilon^2 = \mathrm{tr}(CC^T) = \|C\|^2, \\ &U^T U = V V^T = \mathbf{1}, \end{aligned}$$

where $U = \{u_{kj}\}, \Lambda = \{\lambda_k\}, V = \{v_{ik}\}$. If the value of r is large enough then $C = \mathbf{0}$. This condition will held if $r \geq \min\{m, n\}$. The minimal value of r, for which the condition $A = U\Lambda V^T$ is fair, is equal to rank of the matrix A. Forsythe, G.E. and Moler, C.B. proofed the next theorem.

For any real-valued $(n \times n)$-matrix A there are two real-valued orthogonal $(n \times n)$-matrices U and V such that $U^T AV$ is the diagonal matrix Λ. The matrices U and V can be organized such that the diagonal elements of Λ have the order

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_r > \lambda_{r+1} = ... = \lambda_n = 0,$$

where r — is the rank of A. Particularly, if A is non-degenerate then

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n > 0.$$

The minimization of (2) with the condition (3) is the problem of a two-variable function $\zeta(x, y)$ approximation with a sum of two pair-wise multiplications $\sum_i \alpha_i(x)\beta_i(y)$ of one-variable functions $\alpha_i(x)$ and $\beta_i(y)$. Below we describe a quadratic algorithm to solve this problem.

Find one, then the other the vectors $\mathbf{u}_k, \mathbf{v}_k$ and the singular values λ_k for $k = 1, ..., r$. В качестве этих векторов берутся нормированные значения векторов The normalized vectors $\mathbf{a}_k$ and $\mathbf{b}_k$ are needed to find $\mathbf{u}_k, \mathbf{v}_k$: $\mathbf{u}_k = \frac{\mathbf{a}_k}{\|\mathbf{a}_k\|}$, $\mathbf{v}_k = \frac{\mathbf{b}_k}{\|\mathbf{b}k\|}$. The vectors $\mathbf{a}_k$ и $\mathbf{b}_k$ are found as limits of vector series $\{\mathbf{a}_{k_i}\}$ и $\{\mathbf{b}_{k_i}\}$, respectively $\mathbf{a}_k = \lim(\mathbf{a}_{k_i})$ и $\mathbf{b}_k = \lim(\mathbf{b}_{k_i})$. The singular value λ_k can be found as the multiplication of the norm of the vectors: $\lambda_k = \|\mathbf{a}_k\| \cdot \|\mathbf{b}_k\|$.

The vector $\mathbf{u}_k, \mathbf{v}_k$ searching procedure begins from the choice of the line $\mathbf{b}_{1_1}$ of the matix A which norm is maximal. For $k = 1$ formulas of the vectors $\mathbf{a}_{1_i}$ and $\mathbf{b}_{\mathbf{1}_\mathbf{i}}$ are:

$$\mathbf{a}_{1_i} = \frac{A\mathbf{b}_{1_i}^T}{\mathbf{b}_{1_i}\mathbf{b}_{1_i}^T}, \qquad \mathbf{b}_{1_{i+1}} = \frac{\mathbf{a}_{1_i}^T A}{\mathbf{a}_{1_i}^T \mathbf{a}_{1_i}}, \qquad i = 1, 2, ...$$

To compute $\mathbf{u}_k, \mathbf{v}_k$ where $k = 2, ..., r$ the formulas above are used. However, the matrix A should be replaced with corrected on the k-th step matix $A_{k+1} = A_k - \mathbf{u}_k \lambda_k \mathbf{v}_k$.

Notice the next feature if the Singular Values Decompposituin. Since the matrices U and V are orthogonal, i.e.

$$U^T U = V V^T = I, \tag{4}$$

where I is a $r \times r$ identity matrix, then from (4) one can show that

$$\begin{aligned} AA^T &= U\Lambda V V^T \Lambda U^T = U\Lambda^2 U^T, \\ A^T A &= V^T \Lambda U^T U \Lambda V = V^T \Lambda^2 V. \end{aligned} \tag{5}$$

If one multiply both parts of this equations from the right to U and V^T than

$$\begin{aligned} AA^T U &= U\Lambda^2, \\ A^T A V^T &= V^T \Lambda^2. \end{aligned} \tag{6}$$

From (6) it follows that the matrix U rows are the eigenvectors of the matrix AA^T, while the squares of the singular values $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_r)$ are its eigenvalues (Wilkinson, J.H.). Also the matrix V lines are eigenvectors of the matrix $A^T A$ while the squares of the singular values are is eigenvalues.

FOR FURTHER READING

Forsythe, G.E., Malcolm, M.A., and Moler, C.B. 1977, Computer Methods for Mathematical Computations (Englewood Cliffs, NJ: Prentice-Hall), Chapter 9.

Golub, G.H., and Van Loan, C.F. 1989, Matrix Computations, 2nd ed. (Baltimore: Johns Hopkins University Press), § 8.3 and Chapter 12.

Wilkinson, J.H., and Reinsch, C. 1971, Linear Algebra, vol. II of Handbook for Automatic Computation (New York: Springer-Verlag), Chapter I.10 by G.H. Golub and C. Reinsch.