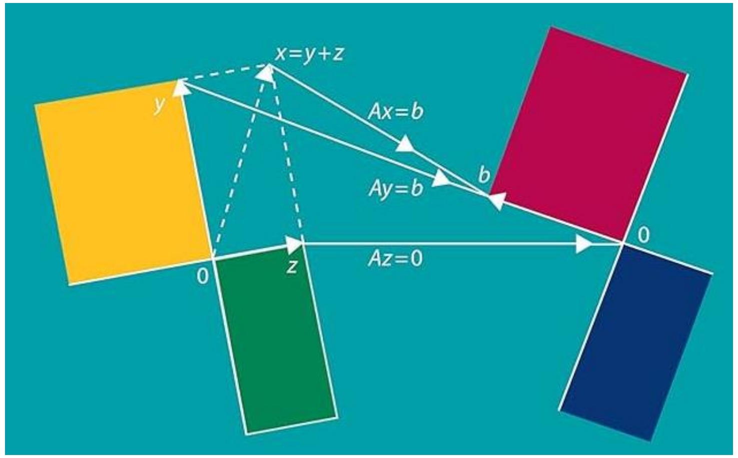
Singular Value Decomposition

(Linear Algebra)

Randil Pushpananda, PhD
rpn@ucsc.cmb.ac.lk



Introduction

Up to this point in the chapter we've dealt exclusively with square matrices. Well, today, we're going to allow rectangular matrices. Is A is an $m \times n$ matrix with $m \neq n$ then the eigenvalue equation

$$A\mathbf{x} = \lambda\mathbf{x}$$

has issues. In particular, the vector $\mathbf{x}$ will have n components, while the vector $A\mathbf{x}$ will have m components (!) and so the equation above won't make sense.

This chapter develops one idea. That idea applies to every matrix, square or rectangular. It is an extension of eigenvectors. But now we need **two sets of orthonormal vectors: input vectors v_1 to v_n and output vectors u_1 to u_m.** This is completely natural for an m by n matrix. The vectors v_1 to v_r are a basis for the row space; u_1 to u_r are a basis for the column space. Then we recover A from r pieces of rank one, with $r = \text{rank}(A)$ and positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ in the diagonal matrix Σ.

$$\boxed{\textbf{SVD} \qquad A = U\Sigma V^{\mathrm{T}} = \sigma_1 u_1 v_1^{\mathrm{T}} + \sigma_2 u_2 v_2^{\mathrm{T}} + \cdots + \sigma_r u_r v_r^{\mathrm{T}}.}$$

The right singular vectors v_i are eigenvectors of $A^{\mathrm{T}}A$. They give bases for the row space and nullspace of A. The left singular vectors u_i are eigenvectors of AA^{T}. They give bases for the column space and left nullspace of A. **Then Av_i equals $\sigma_i u_i$ for $i \leq r$.** The matrix A is diagonalized by these two orthogonal bases: $AV = U\Sigma$.

Introduction

- We need a square matrix.
- Matrices A^TA and AA^T will be square.
- They will also be symmetric.
- Making use of AA^T and A^TA, we'll construct the singular value decomposition of A.

Singular Value Decomposition

Suppose A is an $m \times n$ matrix with rank r. The matrix AA^T will be $m \times m$ and have rank r. The matrix A^TA will be $n \times n$ and also have rank r. Both matrices A^TA and AA^T will be positive semidefinite, and will therefore have r (possibly repeated) positive eigenvalues, and r linearly independent corresponding eigenvectors. As the matrices are symmetric, these eigenvectors will be orthogonal, and we can choose them to be orthonormal.

We call the eigenvectors of A^TA corresponding to its non-zero eigenvalues $\mathbf{v}_1, \ldots, \mathbf{v}_r$. These vectors will be in the row space of A. We call the eigenvectors of AA^T corresponding to its non-zero eigenvalues $\mathbf{u}_1, \ldots, \mathbf{u}_r$. These vectors will be in the column space of A.

Orthogonal

- Two vectors x and y are said to be "orthogonal" if they make an angle of 90 degrees with each other.
- The vectors ***x*** and ***y*** are orthogonal if their dot product is equal to 0.

$$\vec{x} = [x_1, x_2, ..., x_n]$$
$$\vec{y} = [y_1, y_2, ..., y_n]$$

$$\vec{x}.\vec{y} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + ... + x_n y_n$$

$$\vec{x}.\vec{y} = 0$$

Orthonormal

The vectors *x* and *y* are now orthonormal if they both are of unit length as well as orthogonal. Mathematically,

$$\vec{x}.\vec{y} = 0$$
$$||x||_2^2 = 1$$
$$||y||_2^2 = 1$$

Orthonormal Matrix

- ***Q*** is orthonormal if the following conditions are satisfied:

$$||\vec{q_i}||_2^2 = 1 \qquad \forall\ i \in [1, n]$$
$$\vec{q_i}.\vec{q_j} = 0 \qquad \forall\ i, j \in [1, n]\ \&\ i \neq j$$

$$Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix}$$
$$= \begin{bmatrix} \vec{q_1} & \vec{q_2} & \cdots & \vec{q_n} \end{bmatrix}$$

- The magnitude of every column of the orthonormal matrix is 1, and each column is perpendicular to the other.

Now, these vectors have a remarkable relation. Namely,

$$A\mathbf{v}_1 = \sigma_1\mathbf{u}_1, A\mathbf{v}_2 = \sigma_2\mathbf{u}_2, \ldots, A\mathbf{v}_r = \sigma_r\mathbf{u}_r$$

where $\sigma_1, \ldots, \sigma_r$ are positive numbers called the *singular values* of the matrix A.

This relation lets us write

$$A\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r \end{pmatrix}\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}.$$

This gives us a decomposition $AV = U\Sigma$.

Noting that the columns of V are orthonormal we can right multiply both sides of this equality by V^T to get $A = U\Sigma V^T$. This is the singular value decomposition of A.

If we want to we can make V and U square. We just append orthonormal vectors $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ in the nullspace of A to V, and orthonormal vectors $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ in the left-nullspace of A to M. We'll still get $AV = U\Sigma$ and $A = U\Sigma V^T$.

This singular value decomposition has a particularly nice representation if we carry through the multiplication of the matrices:

$$A = U\Sigma V^T = \mathbf{u}_1\sigma_1\mathbf{v}_1 + \cdots + \mathbf{u}_r\sigma_r\mathbf{v}_r^T.$$

Each of these "pieces" has rank 1. If we order the singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$$

then the singular value decomposition gives A in r rank 1 pieces in *order of importance*.