Matrix Decomposition

A matrix decomposition is a factorization of a matrix into some canonical form, that is useful in many scientistic applications including neural networks.

Eigenvalue Decomposition :Let $\mathbf{A}$ be a square $n \times n$ matrix.

- If all eigenvalues of $\mathbf{A}$ are distinct, then $\mathbf{A}$ has n linearly independent eigenvectors.

- Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors of $\mathbf{A}$ and let $\mathbf{T} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$. then

$$\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

  where $\lambda_1, ..., \lambda_n$ are eigenvalues.

- eigenvalues of a symmetric matrix are real

- eigenvectors of a real, symmetric matrix are mutually orthogonal.

Matlab syntax:

$\gg d = eig(A)$

This returns a vector of the eigenvalues of matrix $\mathbf{A}$, i.e. $\lambda_1, ..., \lambda_n$.

$\gg [T, D] = eig(A)$

This produces matrices of eigenvalues (D) and eigenvectors (T) of matrix $\mathbf{A}$.

The QR Factorization: The QR factorization of an $m \times n$ matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{QR}$$

where $\mathbf{Q} \in \Re^{m \times n}$ is orthogonal and $\mathbf{R} \in \Re^{n \times n}$ is upper triangular. Here we assume $m \geq n$. The calculation of the QR factorization is one way of computing an orthonormal basis for a set of vectors.

Gram-Schmidt orthogonalisation: Given an arbitrary set of linearly independent vectors $\mathbf{a}_1$, ..., $\mathbf{a}_n$ in $\mathbf{R}^m$, $(\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n])$ the Gram-Schmidt procedure generates a set of mutually orthogonal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ $(\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_n])$ as follows:

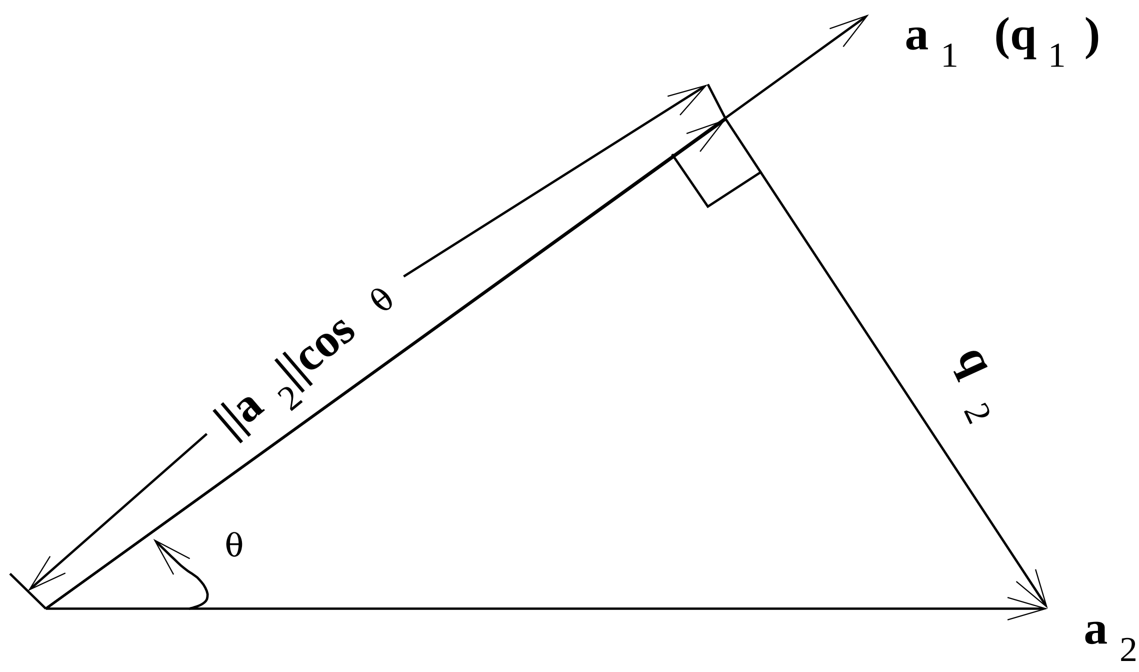
$$\mathbf{q}_1 = \mathbf{a}_1 \tag{1}$$

$$\mathbf{q}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} r_{i,j}\mathbf{q}_i, \quad j = 2, \ldots, n \tag{2}$$

where $r_{i,j} = \frac{\mathbf{q}_i^T \mathbf{a}_j}{\mathbf{q}_i^T \mathbf{q}_i}$. This also produces

$$\mathbf{R} = \begin{bmatrix} 1 & r_{1,2} & r_{1,3} & \cdots & r_{1,n} \\ 0 & 1 & r_{2,3} & \cdots & r_{2,n} \\ 0 & \ldots & \cdots & \ldots & r_{n-1,n} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The Gram-Schmidt orthogonalisation procedure has clear geometric interpretation, as shown in the Figure below.



$\mathbf{q}_2$ is generated by the orthogonal projection of $\mathbf{a}_2$ towards $\mathbf{q}_1$. Note that

$$\cos\theta = \frac{\mathbf{q}_1^T \mathbf{a}_2}{\|\mathbf{q}_1\| \|\mathbf{a}_2\|}$$

and the unit vector in the direction of $\mathbf{a}_1$ (also $\mathbf{q}_1$) is

$$\frac{\mathbf{q}_1}{\|\mathbf{q}_1\|}$$

Hence

$$\mathbf{q}_2 = \mathbf{a}_2 - \|\mathbf{a}_2\| \cos\theta \cdot \frac{\mathbf{q}_1}{\|\mathbf{q}_1\|}$$

$$= \mathbf{a}_2 - \|\mathbf{a}_2\| \frac{\mathbf{q}_1^T \mathbf{a}_2}{\|\mathbf{q}_1\| \|\mathbf{a}_2\|} \cdot \frac{\mathbf{q}_1}{\|\mathbf{q}_1\|}$$

$$= \mathbf{a}_2 - \frac{\mathbf{q}_1^T \mathbf{a}_2}{\|\mathbf{q}_1\|} \cdot \frac{\mathbf{q}_1}{\|\mathbf{q}_1\|}$$

$$= \mathbf{a}_2 - r_{1,2} \mathbf{q}_1$$

Matlab syntax:

$\gg [Q, R] = qr(A)$

Singular Value Decomposition (SVD) :Let $\mathbf{A}$ be a real $m \times n$ matrix, then there exist orthonormal matrices

$$\mathbf{U} = [\mathbf{u}_1, ..., \mathbf{u}_m] \in \Re^{m \times m}$$

and

$$\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_n] \in \Re^{n \times n}$$

such that

$$\mathbf{U}^T \mathbf{A} \mathbf{V} = \text{diag}\{\sigma_1, ..., \sigma_p\} \in \Re^{m \times n}, \quad p = \min\{m, n\}$$

where $\sigma_1 \geq \sigma_2 ... \geq \sigma_p \geq 0$.

The σ_i are the singular values of $\mathbf{A}$, and the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are the ith left singular vectors and the ith right singular vectors respectively.

Matlab syntax:

$\gg [U, S, V] = svd(A)$

Example:

$$\mathbf{A} = \begin{bmatrix} .96 & 1.72 \\ 2.26 & 0.96 \end{bmatrix} = \mathbf{USV}^T$$

$$= \begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}^T$$

Sensitivity Analysis using Singular Value Decomposition (SVD):

Consider the linear system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is a nonsingular n by n matrix and $\mathbf{b}$ is a n by 1 vector.

If $\mathbf{A} = \mathbf{USV}^T$, then

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \sum_{i=1}^{n} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i$$

This shows that if σ_n is small, the large change in $\mathbf{x}$ can be induced by small changes in $\mathbf{A}$ or $\mathbf{b}$.