«Random Processes»

Lecture: Singular Value Decomposition

Andreas Polydoros

Introduction

- SVD is similar to (but not the same as) the eigenvalue-eigenvector factorization of a Hermitian symmetric matrix: $A = E \Lambda E^H$ where:
  - The eigenvalues are in the diagonal matrix Λ.
  - The eigenvector matrix E is orthogonal: $EE^H = I$, because eigenvectors of a symmetric matrix can be chosen to be orthonormal.
  - In general for a general matrix A this is not true.

- If we take out E and E^H above and instead put in Q_1 and Q_2 which are *any two orthogonal matrices* – not necessarily transposes of one another– then the factorization becomes possible for any A.
- The "diagonal" (in general, rectangular) matrix in the middle can be made nonnegative and denoted by Σ.
  - Σ has *positive* entries: $\sigma_1, \ldots, \sigma_r$, which are the **singular values** of A.
  - They fill the first *r* places on the main diagonal of Σ, and *r* is the rank of A.

Definition

Singular Value Decomposition

Any m by n matrix A can be factored into:

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

$$= \text{(orthogonal)(“diagonal”)(orthogonal)}$$

The columns of Q_1 $(m \times m)$ are eigenvectors of AA^H.

The columns of Q_2 $(n \times n)$ are eigenvectors of $A^H A$.

The r singular values on the diagonal of Σ $(m \times n)$ are the square roots of the nonzero eigenvalues of AA^T or $A^T A$ (they are the same).

Examples

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

1. $$A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1]$$

$A^T A$ is 1×1, AA^T is 3×3, they both have eigenvalue 9. The two zero eigenvalues of AA^T leave some freedom for the eigenvectors in columns 2 and 3 of Q_1. The specific choice keep that matrix orthogonal.

2. $$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{\ } \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \quad \text{with} \quad \lambda = 3,1$$

3. $$A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Construction of the SVD

- Any matrix can be factored into $A = Q_1 \Sigma Q_2^T$, with diagonal Σ and orthogonal Q's.
- A symmetric matrix like $A^T A$ has a complete set of orthonormal eigenvectors x_j, which go into the columns of Q_2:

$$A^T A x_j = \lambda_j x_j \text{ with } \quad x_j^T x_j = 1 \qquad \text{and} \quad x_i^T x_j = 0 \quad \text{for} \quad i \neq j$$

  - In the complex case replace A^T with A^H and x^T to x^H.

- Taking the inner product with x_j , we prove easily that all $\lambda_j \geq 0$:

$$x_j^T A^T A x_j = \lambda_j x_j^T x_j \qquad \text{or} \quad \|A x_j\|^2 = \lambda_j$$

- Suppose the eigenvalues $\lambda_1, \ldots, \lambda_r$ are positive, and the remaining $n - r$ λ_j are zero.
  - For the positive ones set $\sigma_j = \sqrt{\lambda_j}$
  - These numbers, $\sigma_1, \ldots, \sigma_r$, will become the singular values on the diagonal of Σ.

- So until now we have the columns of Q_2 (by eigenvectors x_j) and the values σ_j of Σ.

Construction of the SVD

- Calculation of the Q_1 columns:
  - Set: $q_j = Ax_j / \sigma_j$
  - Note that the q_j are unit vectors in $\mathbf{R}^m$ and mutually orthogonal:

$$q_i^T q_j = \frac{x_i^T A^T A x_j}{\sigma_i \sigma_j} = \frac{\lambda_j x_i^T x_j}{\sigma_i \sigma_j} = 0 \quad \text{for} \quad i \neq j$$

  - These *r* orthonormal q's can be extended, by Gram-Schmidt, to a complete orthonormal basis $q_1, \ldots, q_m$. Those are the columns of Q_1.

- The *i,j* entry of the product $Q_1^T \Sigma Q_2$ is $q_i^T A x_j$ (row times matrix times column), and we know those entries:

$$q_i^T A x_j = 0 \quad \text{if} \quad j > r \quad (\text{because} \quad Ax_j = 0)$$

$$q_i^T A x_j = q_i^T \sigma_j q_j \quad \text{if} \quad j \leq r \quad (\text{because} \quad Ax_j = \sigma_j q_j)$$

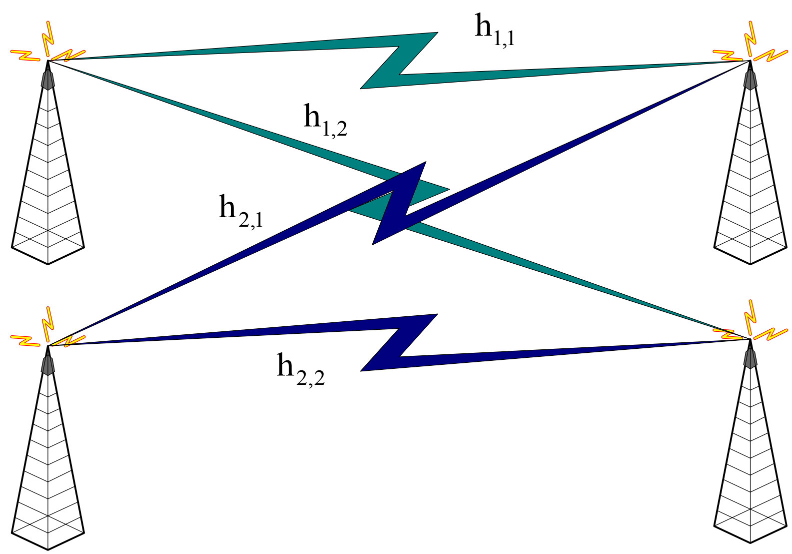
- But $q_i^T q_j = 0$ except when $i = j$. Thus the only nonzeros in the product $\Sigma = Q_1^T A Q_2$ are the first r diagonal entries, which are $\sigma_1, \ldots, \sigma_r$. Since $Q^T = Q^{-1}$ we have $A = Q_1 \Sigma Q_2^T$ as required.

Applications

- Image processing
  - Suppose a satellite sent a 1000×1000 pixel picture to earth.
  - We can code the colors in numbers and send 1.000.000 numbers.
  - With SVD the key is in the singular values (in Σ). Typically some are significant and others extremely small.
  - If we keep 60 and throw away 940 we must send only the corresponding 60 columns of Q_1 and Q_2.
  - The other 940 columns are multiplied in $Q_1 \Sigma Q_2^T$ with small σ's and can be ignored.
  - Then we send 60×2000 = 120.000 numbers instead of 1.000.000.
  - The picture quality will be good or adequate. By adding more singular values the quality will improve.

Applications

MIMO 2x2



$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Transmitted signal from 2 antennas

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \underbrace{\begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix}}_{H} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Received signal from 2 antennas

Applications

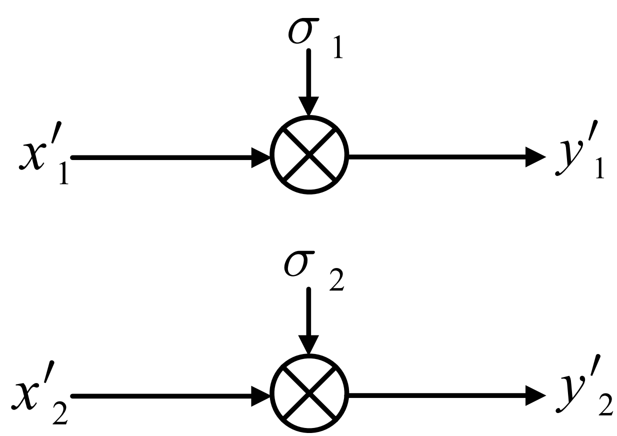
SVD of H: $$H = Q_1 \Sigma Q_2^T = Q_1 \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} Q_2^T$$

The received signals can be transformed as:

$$\underline{y} = Q_1 \Sigma Q_2^T \underline{x}$$

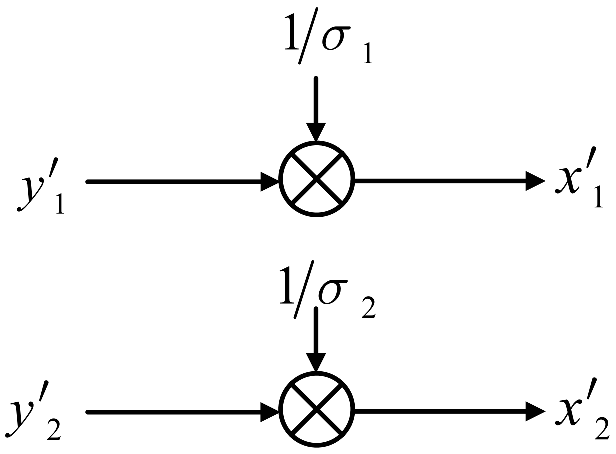
$$\Rightarrow \underbrace{Q_1^T \underline{y}}_{\underline{y}'} = \Sigma \underbrace{Q_2^T \underline{x}}_{\underline{x}'}$$

Equivalent channel:



Applications

Equalization:



The transmitted symbols are:

$$\underline{x} = Q_2 \underline{x}'$$