MODULE 1.2
Singular Value Decomposition (SVD)

by
S.Lakshmivarahan
School of Computer Science
University of Oklahoma
Norman, OK-73019, USA
varahan@ou.edu

What is SVD?

- This module 1.1 contains results relating to the spectral decomposition of square, real symmetric matrices
- This module 1.2 contains analogous results for rectangular matrices, $H \in R^{m \times n}$ called SVD of H
- SVD rests on the spectral decomposition of symmetric matrices $H^T H$ and HH^T are called the Gramian of H

Gramians of H

- Given $H \in R^{m \times n}$, define two related square, symmetric matrices: $H^T H \in R^{n \times n}$ and $HH^T \in R^{m \times m}$ called the Gramians of H
- Assume that H is of full rank, that is,

$$RANK(H) = min(n, m) \tag{1}$$

- From

$$RANK(H^T H) = RANK(H) = RANK(HH^T) \tag{2}$$

it follows that

$$RANK(H^T H) = RANK(HH^T) = min(n, m) \tag{3}$$

- Hence, when $m > n$, $H^T H \in R^{n \times n}$ is non singular and in fact, is SPD. But HH^T is singular and non-negative definite

Spectral decomposition of $H^T H \in R^{n\times n}$ when $m > n$

- Since the smaller Gramian $H^T H$ is an SPD matrix, there exists eigenpairs (λ_i, v_i) $1 \leq i \leq n$ such that

$$(H^T H)V = V\Lambda \tag{4}$$

where $V = [v_1, v_2, \ldots, v_n] \in R^{n\times n}$ and $\Lambda = Diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \in R^{n\times n}$ and $V^T V = VV^T = I_n$

- Also, assume that

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0 \tag{5}$$

- Hence,

$$V^T(H^T H)V = \Lambda \quad \text{and} \quad H^T H = V\Lambda V^T \tag{6}$$

Eigenpair of $HH^T \in R^{m \times m}, m > n$

- Define

$$u_i = \frac{1}{\sqrt{\lambda_i}} Hv_i \in R^m, 1 \leq i \leq n \tag{7}$$

- Then

$$(H^T H)u_i = \frac{1}{\sqrt{\lambda_i}} H(H^T H)v_i = \frac{\lambda_i}{\sqrt{\lambda_i}} Hv_i = \lambda_i u_i \tag{8}$$

- That is, if (λ_i, v_i) is an eigenpair of $(H^T H)$, then (λ_i, u_i) is an eigenpair of HH^T with u_i given by (7)

Spectral decomposition of $HH^T \in R^{m\times m}, m > n$

- Let $U = [u_1, u_2, \ldots, u_n] \in R^{m\times n}$. Then (7) is equivalent to

$$(HH^T)U = U\Lambda \tag{9}$$

- The n non-zero eigenvalues of (HH^T) are the same as the n eigenvalues of $H^T H$. The rest of the (m-n) eigenvalues of HH^T are zero
- The eigenvectors u_i corresponding to the n non-zero eigenvalues of (HH^T) are related to those of $(H^T H)$ through the linear transformation in (7)

- Relation (7) becomes

$$Hv_i = u_i\sqrt{\lambda_i}, \;\; 1 \leq i \leq n \tag{10}$$

- Define

$$U = [u_1, u_2, u_3, \ldots, u_n] \quad \in R^{m\times n}$$

$$\Lambda^{\frac{1}{2}} = Diag(\lambda_1^{\frac{1}{2}}, \lambda_2^{\frac{1}{2}}, \ldots, \lambda_n^{\frac{1}{2}}) \quad \in R^{n\times n}$$

- The n relations in (10) can be written succinctly as

$$HV = U\lambda^{\frac{1}{2}} \quad or \quad H = U\lambda^{\frac{1}{2}}V^T \tag{11}$$

called the SVD of H

Has a sum of rank-1 matrices

- Equation(11) on expanding:

$$H = \begin{bmatrix} u_1, u_2, u_3, \dots, u_n \end{bmatrix} \begin{bmatrix} \lambda_1^{\frac{1}{2}} & 0 & \dots & 0 \\ 0 & \lambda_2^{\frac{1}{2}} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \quad (12)$$

$$= \sum_{i=1}^{n} \lambda_i^{\frac{1}{2}} u_i v_i^T$$

- λ_i 's are the eigenvalues of $(H^T H)$ and are known as the singular values of H
- Hence the name SVD

- Multiplying both sides of (7) on the left by H^T and using (7):

$$H^T u_i = \frac{1}{\sqrt{\lambda_i}}(H^T H)v_i = \frac{1}{\sqrt{\lambda_i}}\lambda_i v_i = \sqrt{\lambda_i} v_i$$

- That is,

$$\begin{aligned} v_i &= \frac{1}{\sqrt{\lambda_i}} H^T u_i \ \ and \\ u_i &= \frac{1}{\sqrt{\lambda_i}} H v_i \end{aligned} \tag{13}$$

are the two defining relations for SVD of H

A Generalization

- Let $\lambda \neq 0, \eta \neq 0$ be such that (λ, η) is an eigenpair of $H^T H$. That is
$$(H^T H)\eta = \lambda\eta \tag{14}$$
- From
$$\lambda(H\eta) = H(H^T H)\eta = (HH^T)(H\eta) \tag{15}$$
it follows that $(\lambda, H\eta)$ is an eigenpair of HH^T
- If $H\eta = 0$, then $(H^T H)\eta = \lambda\eta = 0$ which implies either $\lambda = 0$ or $\eta = 0$ or both zero, which is a contradiction.
- Hence $(\lambda, H\eta)$ is an eigenpair of (HH^T) if (λ, η) is that of $H^T H$

Algebraic and geometric multiplicities of eigenvalues of $H^T H$

- Let λ be an eigenvalue of $(H^T H)$ of algebraic multiplicity, say, m.
- Then, recall that there exists a (non- unique) set of m orthonormal eigenvectors $\{\eta_1, \eta_2, \eta_3, \ldots, \eta_m\}$ such that

$$(H^T H)\eta_i = \lambda \eta_i \quad for \quad 1 \leq i \leq m \tag{16}$$

Algebraic and geometric multiplicities of eigenvalues of HH^T

- Let η_1 and η_2 be two orthogonal eigenvectors of $H^T H$ for the eigenvalue λ of algebraic multiplicity m $= 2$
- Then, $H\eta_1$ and $H\eta_2$ as eigenvectors of (HH^T) are orthogonal
- For
$$(H\eta_1)^T(H\eta_2) = \eta_1^T(H^T H)\eta_2 = \lambda\eta_1^T\eta_2 = 0 \qquad (17)$$

- In view of (15) and (17), the following claim holds:
- Claim: Let H be an $m \times n$ matrix of full rank.
  Then
  (1) The Gramians $H^T H$ and HH^T share the same set of non-zero eigenvalues, and
  (2) λ is an eigenvalue of multiplicity m of $(H^T H)$ with an orthogonal set of eigenvectors $\{\eta_1, \eta_2, \eta_3, \ldots, \eta_m\}$, then λ is also an eigenvalue of multiplicity m of (HH^T) with an orthogonal set of eigenvectors $\{H\eta_1, H\eta_2, \ldots, H\eta_m\}$

Spectral decomposition of $HH^T \in R^{m\times m}, n > m$

- For completeness, we consider the case when $n > m$
- Since (HH^T) is SPD, there exist (λ_i, u_i), $1 \leq i \leq n$ that are eigenpairs of HH^T
- That is,

$$\begin{aligned} (HH^T u_i) = \lambda_i u_i, \quad u_i \in R^n \\ or \quad (HH^T)U = U\lambda \end{aligned} \tag{18}$$

where $U = [u_1, u_2, u_3, \ldots, u_n], U^T U = UU^T = I_m$
$\Lambda = Diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$
where

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \tag{19}$$

Eigenpair of $H^T H \in R^{n \times n}$, n >m

- Define

$$v_i = \frac{1}{\sqrt{\lambda_i}} H^T u_i \in R^n \tag{20}$$

- Then

$$(H^T H) v_i = \frac{1}{\sqrt{\lambda_i}} H^T (H H^T) u_i = \frac{\lambda_i}{\sqrt{\lambda_i}} H^T u_i = \lambda_i v_i \tag{21}$$

- That is, (λ_i, v_i) is an eigenpair of $H^T H$

Eigen decomposition of $H^T H$

- Define

$$V = [v_1, v_2, v_3, \ldots, v_n] \in R^{n \times n}$$

- Then (21) becomes

$$(H^T H)V = V\Lambda, \quad vv^T = I_n \tag{22}$$

- Also the m non-zero eigenvalues of HH^T are those of HH^T and the rest of (n-m) eigenvalues of $H^T H$ are zero.

Dual of (20)

- Multiplying both sides of (20) on the left by H and using (18):

$$Hv_i = \frac{1}{\sqrt{\lambda_i}}(HH^T)u_i = \sqrt{\lambda_i}u_i$$

$$or \quad u_i = \frac{1}{\sqrt{\lambda_i}}Hv_i \tag{23}$$

which is the dual of (20)

A note on our notation

- In this and in all Modules to follow, we use the following convention: $H \in R^{m\times n}$
- Case 1: $m > n$ and $H^T H$ is SPD

$$
\begin{aligned}
(H^T H)V = V\Lambda, \quad V^T V = VV^T = I_n \\
(HH^T)U = U\Lambda, \quad U^T U = I_n, \quad U \in R^{m\times n}
\end{aligned}
\tag{24}
$$

- Case 2: $n > m$ and HH^T is SPD

$$
\begin{aligned}
(HH^T)U = U\Lambda, \quad U^T U = UU^T = I_m \\
(H^T H)V = V\Lambda, \quad V^T V = I_m, \quad V \in R^{n\times m}
\end{aligned}
\tag{25}
$$

A dual characterization of SVD

- Case 1: $m > n$

$$\begin{aligned} u_i &= \frac{1}{\sqrt{\lambda_i}} H v_i \\ v_i &= \frac{1}{\sqrt{\lambda_i}} H^T v_i \end{aligned} \tag{26}$$

- Case 2: $n > m$

$$\begin{aligned} v_i &= \frac{1}{\sqrt{\lambda_i}} H^T u_i \\ u_i &= \frac{1}{\sqrt{\lambda_i}} H v_i \end{aligned} \tag{27}$$