

CSL361 Problem set 7: SVD, Orthogonal projections

October 9, 2009

1 SVD

1. Work out again the *SVD* theorem done in the class:
If A is a real $m \times n$ matrix then there exist orthogonal matrices

$$\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m] \in \mathbb{R}^{m \times m} \text{ and } \mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n] \in \mathbb{R}^{n \times n}$$

such that

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

where

$$\mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_p) \quad p = \min\{m, n\}$$

and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p$.

2. Suppose that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0$. Then show that

(a) $\text{rank}(\mathbf{A}) = r$

(b) $\text{null}(\mathbf{A}) = [\mathbf{v}_{r+1}, \dots, \mathbf{v}_n]$

(c) $\text{range}(\mathbf{A}) = [\mathbf{u}_1, \dots, \mathbf{u}_r]$

(d) If $\mathbf{U}_r = \mathbf{U}(:, 1:r)$, $\mathbf{\Sigma}_r = \mathbf{\Sigma}(1:r, 1:r)$ and $\mathbf{V}_r = \mathbf{V}(:, 1:r)$, then

$$\mathbf{A}_r = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

3. Show that

(a) $\|\mathbf{A}\|_F^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_p^2$

(b) $\|\mathbf{A}\|_2^2 = \sigma_1$

4. Let the *SVD* of \mathbf{A} be as above. If $k < r = \text{rank}(\mathbf{A})$ and

$$\mathbf{A}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

then show that

$$\min_{\text{rank}(\mathbf{B}=k)} \|\mathbf{A} - \mathbf{B}\|_2^2 = \|\mathbf{A} - \mathbf{A}_k\|_2^2 = \sigma_1$$

5. Show (without using condition numbers) that if \mathbf{A} is square ($n \times n$) and $\sigma_n > 0$ is small, then solving $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ is unstable.
6. Show that in the $\|\cdot\|_2$ norm

$$\text{cond}(\mathbf{A}) = \frac{\sigma_1}{\sigma_n}$$

2 Least-squares

1. Consider the least-squares problem:
Find the least-squares solution to the $m \times n$ set of equations $\mathbf{Ax} = \mathbf{b}$, where $m > n$ and $\text{rank}(\mathbf{A}) = n$
 Show that the following constitute a solution:
- (a) Find the *SVD* $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$
 - (b) Set $\mathbf{b}' = \mathbf{U}^T\mathbf{b}$
 - (c) Find the vector \mathbf{y} defined by $y_i = b'_i/\sigma_i$, where σ_i is the i^{th} diagonal entry of $\mathbf{\Sigma}$
 - (d) The solution is $\mathbf{x} = \mathbf{V}\mathbf{y}$
2. Consider the least-squares problem:
Find the general least-squares solution to the $m \times n$ set of equations $\mathbf{Ax} = \mathbf{b}$, where $m > n$ and $\text{rank}(\mathbf{A}) = r < n$
 Show that the following constitute a solution:
- (a) Find the *SVD* $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$
 - (b) Set $\mathbf{b}' = \mathbf{U}^T\mathbf{b}$
 - (c) Find the vector \mathbf{y} defined by $y_i = b'_i/\sigma_i$, for $i = 1, \dots, r$, and $y_i = 0$ otherwise.

- (d) The solution \mathbf{x} of minimum norm $\|\mathbf{x}\|$ is $\mathbf{V}\mathbf{y}$
 (e) The general solution is

$$\mathbf{x} = \mathbf{V}\mathbf{y} + \lambda_{r+1}\mathbf{v}_{r+1} + \dots + \lambda_n\mathbf{v}_n$$

where $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ are the last $n - r$ columns of \mathbf{V} .

3. Given a square diagonal matrix Σ , its *pseudo-inverse* is defined as the diagonal matrix Σ^+ such that

$$\sigma_{ii}^+ = \begin{cases} 0 & \text{if } \sigma_{ii} = 0 \\ \Sigma_{ii}^{-1} & \text{otherwise} \end{cases}$$

For an $m \times n$ matrix A with $m \geq n$ with $A = U\Sigma V^t$, its *pseudo-inverse* is defined as

$$A^+ = V\Sigma^+U^t$$

Show that the least-squares solution to an $m \times n$ system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ of rank n is given by $\mathbf{A}^+\mathbf{b}$ (pseudo-inverse). In the case of a deficient-rank system, $\mathbf{x} = \mathbf{A}^+\mathbf{b}$ is the solution that minimizes $\|\mathbf{x}\|$.

4. Show that if \mathbf{A} is an $m \times n$ matrix of rank n , then $\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ and, in general, a least-squares solution can be obtained by solving the *normal equations*

$$(\mathbf{A}^T\mathbf{A})\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

5. **Weighted least-squares:** Let \mathbf{C} be a positive definite matrix. Then the \mathbf{C} -norm is defined as $\|\mathbf{a}\|_{\mathbf{C}} = (\mathbf{a}^T\mathbf{C}\mathbf{a})^{1/2}$. The *weighted least-squares* problem is one of minimizing $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_{\mathbf{C}}$. The most common weighting is when \mathbf{C} is diagonal. Show that weighted least-squares solution can be obtained by solving:

$$(\mathbf{A}^T\mathbf{C}\mathbf{A})\mathbf{x} = \mathbf{A}^T\mathbf{C}\mathbf{b}$$

6. Consider the constrained least squares problem:
 Given \mathbf{A} of size $n \times n$, find \mathbf{x} that minimizes $\|\mathbf{A}\mathbf{x}\|$ subject to $\|\mathbf{x}\| = 1$. Show that the solution is given by the last column of \mathbf{V} where $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ is the *SVD* of \mathbf{A} .
7. Consider the following constrained least-squares problem: Given an $m \times n$ matrix \mathbf{A} with $m \geq n$, find the vector \mathbf{x} that minimizes $\|\mathbf{A}\mathbf{x}\|$ subject to $\|\mathbf{x}\| = 1$ and $\mathbf{C}\mathbf{x} = \mathbf{0}$. Show that a solution is given as:

- (a) If \mathbf{C} has fewer rows than columns, then add $\mathbf{0}$ rows to \mathbf{C} to make it square. Compute the *SVD* $\mathbf{C} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$. Let \mathbf{C}^\perp be the matrix obtained from \mathbf{V} after deleting the first r columns where r is the number of non-zero entries in $\mathbf{\Sigma}$.
 - (b) Find the solution to minimization of $\|\mathbf{A}\mathbf{C}^\perp\mathbf{x}'\|$ subject to $\|\mathbf{x}'\| = 1$.
 - (c) The solution is obtained as $\mathbf{x} = \mathbf{C}^\perp\mathbf{x}'$.
8. Consider the following constrained least-squares problem: *Given an $m \times n$ matrix \mathbf{A} with $m \geq n$, find the vector \mathbf{x} that minimizes $\|\mathbf{A}\mathbf{x}\|$ subject to $\|\mathbf{x}\| = 1$ and $\mathbf{x} \in \text{range}(\mathbf{G})$.*
Show that a solution is given as:
- (a) Compute the *SVD* $\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$. Let
 - (b) Let \mathbf{U}' be the matrix of first r columns of \mathbf{G} where $r = \text{rank}(\mathbf{G})$.
 - (c) Find the solution to minimization of $\|\mathbf{A}\mathbf{U}'\mathbf{x}'\|$ subject to $\|\mathbf{x}'\| = 1$.
 - (d) The solution is obtained as $\mathbf{x} = \mathbf{U}'\mathbf{x}'$.

3 Orthogonal Projections

Let $S \subseteq \mathbb{R}^n$ be a subspace. $\mathbf{P} \in \mathbb{R}^{n \times n}$ is the *orthogonal projection* onto S if $\text{range}(\mathbf{P}) = S$, $\mathbf{P}^2 = \mathbf{P}$ and $\mathbf{P}^T = \mathbf{P}$.

1. Show the following:
 - (a) If $\mathbf{x} \in \mathbb{R}^n$ and \mathbf{P} is an orthogonal projection on to S ($\mathbf{P}\mathbf{x} \in S$), then $(\mathbf{I} - \mathbf{P})$ is an orthogonal projection onto S^\perp ($(\mathbf{I} - \mathbf{P})\mathbf{x} \in S^\perp$) where S^\perp is the orthogonal complement of S .
 - (b) The *orthogonal projection* onto a subspace is unique.
 - (c) If $\mathbf{v} \in \mathbb{R}^n$, then $\mathbf{P} = \mathbf{v}\mathbf{v}^T / \mathbf{v}^T\mathbf{v}$ is the orthogonal projection onto $S = \text{span}\{\mathbf{v}\}$.
 - (d) If the columns of $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_k]$ are an orthonormal basis for S , then $\mathbf{V}\mathbf{V}^T$ is the unique orthonormal projection onto S .
2. Suppose that the *SVD* of \mathbf{A} is $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ and $\text{rank}(\mathbf{A}) = r$. If we have the \mathbf{U} and \mathbf{V} partitionings:

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_r & \tilde{\mathbf{U}}_r \\ r & m-r \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_r & \tilde{\mathbf{V}}_r \\ r & n-r \end{bmatrix}$$

Then, show that

- (a) $\mathbf{V}_r \mathbf{V}_r^T = \text{projection onto } \text{null}(\mathbf{A})^\perp = \text{range}(\mathbf{A}^T)$
- (b) $\tilde{\mathbf{V}}_r \tilde{\mathbf{V}}_r^T = \text{projection onto } \text{null}(\mathbf{A})$
- (c) $\mathbf{U}_r \mathbf{U}_r^T = \text{projection onto } \text{range}(\mathbf{A})$
- (d) $\tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T = \text{projection onto } \text{range}(\mathbf{A})^\perp = \text{null}(\mathbf{A}^T)$

3. Let $\mathbf{v} \in \mathbb{R}^n$ be non-zero. The $n \times n$ matrix

$$\mathbf{P} = \mathbf{I} - 2 \frac{\mathbf{v} \mathbf{v}^T}{\mathbf{v}^T \mathbf{v}}$$

is the *Householder reflection* discussed in class. Show that:

- (a) \mathbf{P} is an orthogonal projection.
- (b) When a vector $\mathbf{v} \in \mathbb{R}^n$ is multiplied by \mathbf{P} , it is reflected in the hyperplane $\text{span}\{\mathbf{v}\}^\perp$.
- (c) If

$$\mathbf{v} = \mathbf{x} \pm \|\mathbf{x}\|_2 \mathbf{e}_1$$

Then,

$$\mathbf{P} \mathbf{x} = \mp \|\mathbf{x}\|_2 \mathbf{e}_1 \in \text{span}\{\mathbf{e}_1\}$$