Problem set 5: SVD, Orthogonal projections, etc.

February 21, 2017

1 SVD

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

   $$\mathbf{U} = [\mathbf{u_1}, \ldots, \mathbf{u_m}] \in \mathbb{R}^{m \times m} \text{ and } \mathbf{V} = [\mathbf{v_1}, \ldots, \mathbf{v_n}] \in \mathbb{R}^{n \times n}$$

   such that

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

   where

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

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

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

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

   (b) $null(\mathbf{A}) = [\mathbf{v_{r+1}}, \ldots, \mathbf{v_n}]$

   (c) $range(\mathbf{A}) = [\mathbf{u_1}, \ldots, \mathbf{u_r}]$

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

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

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

$$\mathbf{A_k} = \sum_{i=1}^{k} \sigma_i \mathbf{u_i} \mathbf{v_i}^T$$

then show that

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

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 $\| \, \|_2$ norm

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

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 $rank(\mathbf{A}) = n$
Show that the following constitute a solution:
   (a) Find the *SVD* $\mathbf{A} = \mathbf{U\Sigma V}^T$
   (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 $\mathbf{x} = \mathbf{Vy}$

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 $rank(\mathbf{A}) = r < n$
Show that the following constitute a solution:
   (a) Find the *SVD* $\mathbf{A} = \mathbf{U\Sigma V}^T$
   (b) Set $\mathbf{b'} = \mathbf{U}^T\mathbf{b}$
   (c) Find the vector $\mathbf{y}$ defined by $y_i = b'_i/\sigma_i$, for $i = 1, \ldots, r$, and $y_i = 0$ otherwise.
   (d) The solution $\mathbf{x}$ of minimum norm $\|\mathbf{x}\|$ is $\mathbf{Vy}$

(e) The general solution is

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

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

3. Show that the least-squares solution to an $m \times n$ system of equations $\mathbf{Ax} = \mathbf{b}$ 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 $\|\mathbf{Ax} - \mathbf{b}\|_{\mathbf{C}}$. The most common weighting is when $\mathbf{C}$ is diagonal. Show that weigthed least-sqaures 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* $\|\mathbf{Ax}\|$ *subject to* $\|\mathbf{x}\| = 1$.
Show that the solution is given by the last column of $\mathbf{V}$ where $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ 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* $\|\mathbf{Ax}\|$ *subject to* $\|\mathbf{x}\| = 1$ *and* $\mathbf{Cx} = \mathbf{0}$.
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 $\|\mathbf{Ax}\|$ subject to $\|\mathbf{x}\| = 1$ and $\mathbf{x} \in range(\mathbf{G})$.*
   Show that a solution is given as:
   (a) Compute the *SVD* $\mathbf{G} = \mathbf{U\Sigma V}^T$. Let
   (b) Let $\mathbf{U}'$ be the matrix of first r columns of $\mathbf{G}$ where $r = rank(\mathbf{G})$.
   (c) Find the solution to minimization of $\|\mathbf{AU}'\mathbf{x}'\|$ 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 $range(\mathbf{P}) = S$, $\mathbf{P}^2 = \mathbf{P}$ 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 ($\mathbf{Px} \in S$), 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 $S = span\{\mathbf{v}\}$.
   (d) If the columns of $\mathbf{V} = [\mathbf{v_1}, \ldots, \mathbf{v_k}]$ are an orthonormal basis for S, then $\mathbf{VV}^T$ is the unique orthonormal projection onto S.

2. Suppose that the *SVD* of $\mathbf{A}$ is $\mathbf{A} = \mathbf{U\Sigma V}^T$ and $rank(\mathbf{A}) = r$. If we have the $\mathbf{U}$ and $\mathbf{V}$ partitionings:

$$\mathbf{U} = [\ \underset{r}{\mathbf{U_r}} \quad \underset{m-r}{\tilde{\mathbf{U}}_\mathbf{r}}\ ] \qquad \mathbf{V} = [\ \underset{r}{\mathbf{V_r}} \quad \underset{n-r}{\tilde{\mathbf{V}}_\mathbf{r}}\ ]$$

   Then, show that
   (a) $\mathbf{V_r}\mathbf{V_r}^T$ = projection onto $null(\mathbf{A})^\perp = range(\mathbf{A}^T)$
   (b) $\tilde{\mathbf{V}}_\mathbf{r}\tilde{\mathbf{V}}_\mathbf{r}^T$ = projection onto $null(\mathbf{A})$
   (c) $\mathbf{U_r}\mathbf{U_r}^T$ = projection onto $range(\mathbf{A})$
   (d) $\tilde{\mathbf{U}}_\mathbf{r}\tilde{\mathbf{U}}_\mathbf{r}^T$ = projection onto $range(\mathbf{A})^\perp = null(\mathbf{A}^T)$

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*. 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 $span\{\mathbf{v}\}^{\perp}$.

(c) If

$$\mathbf{v} = \mathbf{x} \pm \|\mathbf{x}\|_2\mathbf{e_1}$$

Then,

$$\mathbf{P}\mathbf{x} = \mp\|\mathbf{x}\|_2\mathbf{e_1} \in span\{\mathbf{e_1}\}$$