Optimization and Computational Linear Algebra – Brett Bernstein

Recitation 7

It may be helpful if after recitation you try to re-solve these problems by yourself, and use them as additional study problems for the class.

1. Let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times m}$ and $C \in \mathbb{R}^{n\times n}$. Prove that if B and C are invertible then $\operatorname{rank}(A) = \operatorname{rank}(BAC)$. What does this say about the spectral decomposition?

   Solution. By the homework we know that $\operatorname{rank}(AC) \leq \operatorname{rank}(A)$. Let $v \in \operatorname{im}(A)$ so that $v = Ax$ for some $x \in \mathbb{R}^n$. Then $v = AC(C^{-1}x)$ so $v \in \operatorname{im}(AC)$. Thus $\operatorname{rank}(AC) = \operatorname{rank}(A)$. For the other side, note that

   $$\operatorname{rank}(BA) = \operatorname{rank}(A^T B^T) = \operatorname{rank}(A^T) = \operatorname{rank}(A)$$

   by applying the previous argument, and noting that B^T is invertible (since $(B^{-1})^T B^T = (BB^{-1})^T = I^T$). Thus

   $$\operatorname{rank}(BAC) = \operatorname{rank}((BA)C) = \operatorname{rank}(BA) = \operatorname{rank}(A).$$

   This proves that the rank of a symmetric matrix is equal to its number of non-zero eigenvalues. To see this suppose the spectral decomposition of M is given by $M = V\Lambda V^T$. Then we have

   $$\operatorname{rank}(M) = \operatorname{rank}(V\Lambda V^T) = \operatorname{rank}(\Lambda)$$

   since V, V^T are orthogonal and thus invertible.

2. Suppose $D \in \mathbb{R}^{n\times n}$ is diagonal. Give a vector $v \in \mathbb{R}^n$ with $\|v\| = 1$ such that $\|Dv\|$ maximized.

   Solution. Note that

   $$\|Dv\|^2 = \sum_{i=1}^n (D_{ii}v_i)^2 \leq \left(\max_i D_{ii}^2\right) \sum_{i=1}^n v_i^2 = \max_i D_{ii}^2.$$

   Thus we can choose $v = e_j$ where $|D_{jj}|$ is the largest absolute diagonal entry of D.

3. Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric. Give a vector v with $\|v\| = 1$ such that $\|Av\|$ maximized.

   Solution. By the spectral theorem we have $A = U\Lambda U^T$ where Λ is diagonal and U is orthogonal. Write v as

   $$v = \alpha_1 u_1 + \cdots + \alpha_n u_n$$

where $u_1, \ldots, u_n$ are the columns of U. Then we have

$$\|Av\|^2 = \|U\Lambda U^T \sum_{i=1}^n \alpha_i u_i\|^2 = \|\sum_{i=1}^n \alpha_i \lambda_i u_i\|^2 = \sum_{i=1}^n \alpha_i^2 \lambda_i^2 \leq \max_i \lambda_i^2 \sum_{i=1}^n \alpha_i^2 = \max_i \lambda_i^2$$

where $\lambda_i = \Lambda_{ii}$. Thus we can choose $v = u_j$ where $|\lambda_j|$ is the largest absolute eigenvalue.

4. Suppose $A \in \mathbb{R}^{m \times n}$. Give a vector w with $\|w\| = 1$ such that $\|Aw\|$ maximized.

   Solution. See extra credit 7.5 on the homework.

5. Let $A \in \mathbb{R}^{n \times n}$ have eigenvalue λ. Prove that

   $$E_\lambda = \{v \in \mathbb{R}^n : Av = \lambda v\}$$

   is a subspace of $\mathbb{R}^n$ (called the eigenspace of A corresponding to λ).

   Solution.

   - $A0 = 0 = \lambda \cdot 0$ so $0 \in E_\lambda$.
   - If $v, w \in E_\lambda$ then

     $$A(v + w) = Av + Aw = \lambda v + \lambda w = \lambda(v + w)$$

     proving $v + w \in E_\lambda$.
   - If $v \in E_\lambda$ and $c \in \mathbb{R}$ then

     $$A(cv) = cAv = c\lambda v = \lambda(cv)$$

     proving $cv \in E_\lambda$.

6. Let $A \in \mathbb{R}^{n \times n}$ have eigenvalue λ. How would you find a non-zero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$?

   Solution. Solve the linear system $(A - \lambda I)v = 0$.

7. Let $A \in \mathbb{R}^{m \times n}$ and let $k = \min(m, n)$. Show there are orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that $A = U\Sigma V^T$ where $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal (rectangular) matrix with non-negative entries (in other words, $\Sigma_{ij} = 0$ if $i \neq j$). The diagonal entries are labeled $\sigma_1 = \Sigma_{11}, \ldots, \sigma_k = \Sigma_{kk}$ and are ordered so that $\sigma_1 \geq \cdots \geq \sigma_k$. This is called the singular value decomposition (SVD) of A, the values $\sigma_1, \ldots, \sigma_k$ are called the singular values of A, the columns of U are called the left singular vectors of A, and the columns of V are called the right singular vectors of A.

   Solution.

Proof. Note that $A^TA \in \mathbb{R}^{n\times n}$ is symmetric, so we can apply the spectral theorem to obtain

$$A^TA = V\Lambda V^T$$

where $V \in \mathbb{R}^{n\times n}$ is orthogonal and $\Lambda \in \mathbb{R}^{n\times n}$ is diagonal. Suppose A^TA has rank r and we order the columns of V so that $\lambda_{r+1} = 0, \dots, \lambda_n = 0$. Let $w_1, \dots, w_n$ denote the n columns of AV. We claim that $w_1, \dots, w_n$ are orthogonal (but not necessarily orthonormal). To see this note that

$$V^TA^TAV = \Lambda,$$

which is diagonal. This proves the claim since $\Lambda_{ij} = w_i^T w_j$. Note also that $w_{r+1}, \dots, w_n$ are zero, since their squared lengths are given by $\lambda_{r+1}, \dots, \lambda_n$, which are assumed to be zero. Let $u_i = w_i/\|w_i\|$ for $i = 1, \dots, r$, and extend with $m - r$ new vectors to form an orthonormal basis

$$u_1, \dots, u_r, u_{r+1}, \dots, u_m.$$

Then we have

$$AV = U\Sigma$$

where $U \in \mathbb{R}^{m\times m}$ has u_i as its ith column and $\Sigma \in \mathbb{R}^{m\times n}$ is given by

$$\Sigma = \begin{bmatrix} \|w_1\| & & & & \\ & \ddots & & & \\ & & \|w_r\| & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix},$$

with zeros on the off-diagonal. Thus $A = U\Sigma V^T$ as required. $\square$

8. Let $A \in \mathbb{R}^{m\times n}$. Give a method for computing $\text{rank}(A)$ using the SVD of A.

   Solution. Writing $A = U\Sigma V^T$ we can simply count the number of non-zero entries in Σ since U, V are invertible.

9. Explain the following statement: For any $A \in \mathbb{R}^{m\times n}$, the set $\{Ax : \|x\| = 1\}$ is an ellipsoid. In other words, the image of the sphere under a linear transformation is always an ellipsoid.

   Solution. Using the SVD write $A = U\Sigma V^T$. V^T is orthogonal, so it preserves lengths and maps the sphere $\{x : \|x\| = 1\}$ to itself. Then Σ stretches the sphere along each axis creating an ellipsoid. Finally U is orthogonal, so it rotates the ellipsoid.