Optimization Prof. Carlson

Homework 5

1. Show that a real symmetric matrix $\mathbf{A}$ is negative definite if and only if $(-1)^k \Delta_k > 0$ for $k = 1, \dots, N$.

Solution: The matrix $\mathbf{A}$ is negative definite if and only if $\mathbf{B} = -\mathbf{I_N}\mathbf{A}$ is positive definite. ($\mathbf{I_N}$ is the $N \times N$ identity.) We've already seen that a real symmetric matrix $\mathbf{B}$ is positive definite if and only $\Delta_k > 0$ for $k = 1, \dots, N$. Thus $\mathbf{A}$ is negative definite if and only if

$$\Delta_k(\mathbf{B}) = \det(-\mathbf{I_k}) \det(\mathbf{A_k}) = (-1)^k \Delta_k(\mathbf{A}) > 0,$$

as desired.

2. Show that if a vector X of norm 1 maximizes the quadratic form

$$X \bullet \mathbf{A}X,$$

for a real symmetric matrix $\mathbf{A}$, then this vector must be an eigenvector corresponding to the largest eigenvalue of $\mathbf{A}$.

Solution: Let $E_1, \dots, E_N$ be an orthonormal basis of eigenvectors for $\mathbf{A}$, with eigenvalues λ_k. Suppose that $\lambda_1 \geq \dots \geq \lambda_N$. Write

$$X = \sum_{k=1}^{N} c_k E_k, \quad \sum c_k^2 = 1.$$

By the orthonormality of the vectors E_k,

$$X \bullet \mathbf{A}X = \sum_{k=1}^{N} c_k E_k \bullet \sum_{k=1}^{N} c_k \lambda_k E_k = \sum_{k=1}^{N} \lambda_k c_k^2 \leq \lambda_1 \sum_{k=1}^{N} c_k^2 = \lambda_1.$$

Also note that the inequality is strict if $|c_k| > 0$ for any k with $\lambda_k < \lambda_1$.

3. Prove the following theorem. Hint: Write X as a linear combination of eigenvectors, and see what happens when you multiply by $\mathbf{A}^n$.

Theorem: *Suppose the real symmetric matrix $\mathbf{A}$ is positive semidefinite. If the vector X is not orthogonal to the eigenspace corresponding to the largest eigenvalue $\lambda_1 > 0$, then the sequence*

$$X_n = \frac{\mathbf{A}^n X}{\|\mathbf{A}^n X\|}$$

converges to an eigenvector of $\mathbf{A}$ with eigenvalue λ_1.

Solution: Let $E_1, \dots, E_N$ be an orthonormal basis of eigenvectors for $\mathbf{A}$, with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_N \geq 0$.

Write

$$X = \sum_{k=1}^{N} c_k E_k, \quad \sum c_k^2 = 1.$$

We have

$$\mathbf{A}^n X = \sum_{k=1}^{N} c_k \lambda_k^n E_k = \lambda_1^n \sum_{k=1}^{N} c_k (\frac{\lambda_k}{\lambda_1})^n E_k$$

and

$$\|\mathbf{A}^n X\| = \lambda_1^n [\sum_{k=1}^{N} c_k^2 (\frac{\lambda_k}{\lambda_1})^{2n}]^{1/2}.$$

If $0 \leq \lambda_k < \lambda_1$ then

$$(\frac{\lambda_k}{\lambda_1})^n \to 0.$$

Thus

$$\lim_n X_n = \lim_n \frac{\mathbf{A}^n X}{\|\mathbf{A}^n X\|} = \frac{\sum_{\lambda_k = \lambda_1} c_k E_k}{[\sum_{\lambda_k = \lambda_1} c_k^2]^{1/2}},$$

which is an eigenvector of $\mathbf{A}$ with eigenvalue λ_1.

Since the vector X is not orthogonal to the eigenspace corresponding to the largest eigenvalue $\lambda_1 > 0$, at least one of the c_k with $\lambda_k = \lambda_1$ is not 0.

4. Find an orthonormal basis of eigenvectors for

$$\mathbf{A} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

Diagonalize $\mathbf{A}$. Compute $\mathbf{A}^{10}$.

Solution: An orthonormal basis of eigenvectors is

$$E_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad E_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

If

$$U = \begin{pmatrix} E_1 & E_2 \end{pmatrix}, \quad U^{-1} = U^T = \begin{pmatrix} E_1^T \\ E_2^T \end{pmatrix}$$

then

$$U^{-1} \mathbf{A} U = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} = D.$$

Thus
$$\mathbf{A}^{10} = (UDU^{-1})^{10} = U \begin{pmatrix} 2^{10} & 0 \\ 0 & 4^{10} \end{pmatrix} U^{-1}.$$

5. What happens in problem 3 if $\mathbf{A}$ is real symmetric, but not positive semidefinite?

Solution: Let the eigenvalues be ordered $\lambda_1 \geq \cdots \geq \lambda_N$. In this case λ_N may be negative. If $\lambda_1 > |\lambda_N|$ the conclusion is the same, essentially by the same argument. If $\lambda_1 < |\lambda_N|$ and $c_N \neq 0$, the conclusion is the same for the sequences X_{2n} and X_{2n+1}, but the subsequential limits differ by sign.

If $\lambda_1 = |\lambda_N|$ the subsequence X_{2n} may have a limit which is a linear combination of eigenvectors for λ_1 and λ_N.