MAPL 600 / CMSC 760 Fall 2007
Take-Home Exam 2
Partial answers

1a. (10) Let $\mathbf{T}_m$ be a symmetric tridiagonal matrix of size $m \times m$ and let $\mathbf{T}_{m-1}$ be the $(m-1) \times (m-1)$ matrix formed from its first $m-1$ rows and columns. Denote the eigenvalues of $\mathbf{T}_{m-1}$ by $\tau_1 \geq \ldots \geq \tau_{m-1}$, and denote the eigenvalues of $\mathbf{T}_m$ by $\lambda_1 \geq \ldots \geq \lambda_m$. Prove that for $j = 1, \ldots, m-1$,

$$\lambda_j \geq \tau_j \geq \lambda_{j+1}.$$

Note: The result is true for general symmetric matrices, but we only need it for the tridiagonal case. Hint: You may use without proof the Courant-Fischer Minimax Theorem.

Answer: Suppose

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ 0 \end{bmatrix}$$

where $\mathbf{w}_1$ is $(m-1) \times 1$ and $\mathbf{w}$ is $m \times 1$. Then observe that

$$\mathbf{w}^T \mathbf{T}_m \mathbf{w} = \mathbf{w}_1^T \mathbf{T}_{m-1} \mathbf{w}_1.$$

Let W_k^n denote a subspace of $\mathcal{R}^n$ of dimension k. Let $\bar{W}_k^n$ denote a subspace of $\mathcal{R}^n$ of dimension k that includes the last column of the identity matrix. Then

$$\begin{aligned}
\tau_j &= \min_{W_{j-1}^{m-1}} \max_{\mathbf{w}_1 \perp W_{j-1}^{m-1}} \frac{\mathbf{w}_1^T \mathbf{T}_{m-1} \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \\
&= \min_{\bar{W}_j^m} \max_{\mathbf{w} \perp \bar{W}_j^m} \frac{\mathbf{w}^T \mathbf{T}_m \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \\
&\geq \min_{W_j^m} \max_{\mathbf{w} \perp W_j^m} \frac{\mathbf{w}^T \mathbf{T}_m \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \\
&= \lambda_{j+1}.
\end{aligned}$$

(The first line follows from the minimax theorem, the second from the observation above, the third from expanding the set of subspaces we minimize over, and the fourth from the minimax theorem.)

Similarly,

$$\begin{aligned}
\tau_j &= \max_{W_{m-j-1}^{m-1}} \min_{\mathbf{w}_1 \perp W_{m-j-1}^{m-1}} \frac{\mathbf{w}_1^T \mathbf{T}_{m-1} \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1} \\
&= \max_{\bar{W}_{m-j}^m} \min_{\mathbf{w} \perp \bar{W}_{m-j}^m} \frac{\mathbf{w}^T \mathbf{T}_m \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \\
&\leq \max_{W_{m-j}^m} \min_{\mathbf{w} \perp W_{m-j}^m} \frac{\mathbf{w}^T \mathbf{T}_m \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \\
&= \lambda_j.
\end{aligned}$$

Note that we have not used the tridiagonal structure of $\mathbf{T}_m$, so the result holds for general symmetric matrices.

1b. (10) This leads to an algorithm for approximating some of the eigenvalues of a large sparse symmetric matrix $\mathbf{A}$: run Arnoldi and use the eigenvalues of $\mathbf{T}_m$ as approximations to some eigenvalues of $\mathbf{A}$. Prove (using the Arnoldi relation and about 2 lines of writing) that $\mathbf{T}_n$ is similar to the $n \times n$ matrix $\mathbf{A}$, and show how to use the eigenvectors of $\mathbf{T}_m$ to form approximations to some eigenvectors of $\mathbf{A}$.

Answer: The Arnoldi relation says

$$\mathbf{A}\mathbf{V}_n = \mathbf{V}_n\mathbf{T}_n,$$

where $\mathbf{V}_n$ is an orthogonal matrix and therefore invertible. Therefore,

$$\mathbf{V}_n^{-1}\mathbf{A}\mathbf{V}_n = \mathbf{T}_n,$$

so $\mathbf{T}_n$ is a similarity tranform of $\mathbf{A}$ and therefore the two matrices have the same eigenvalues. Further, if $\mathbf{T}_n\mathbf{z} = \lambda\mathbf{z}$ (so that $\mathbf{z}$ is an eigenvector of $\mathbf{T}_n$ corresponding to eigenvalue λ), then, starting with the Arnoldi relation we have

$$\mathbf{A}\mathbf{V}_n\mathbf{z} = \mathbf{V}_n\mathbf{T}_n\mathbf{z} = \lambda\mathbf{V}_n\mathbf{z},$$

so $\mathbf{V}_n\mathbf{z}$ is an eigenvector of $\mathbf{A}$ corresponding to eigenvalue $\boldsymbol{\lambda}$.

So our algorithm involves running Arnoldi for m steps, finding the eigenvalues and eigenvectors of $\mathbf{T}_m$, and using the eigenvalues and $\mathbf{V}_m$ times the eigenvectors as approximations to the eigenvalues and eigenvectors of $\mathbf{A}$. We know that our estimated eigenvalues will always lie in the interval between the largest and smallest eigenvalues of $\mathbf{A}$ (by 1a).

1c. (10) Implement your algorithm to approximate some of the eigenvalues and eigenvectors of a symmetric positive definite $\mathbf{A}$, by modifying `cg.m` (available on the website) or by writing your own program. Use Matlab's `eig` to find the eigenvalues and eigenvectors of $\mathbf{T}_m$ (although faster algorithms exist, which take advantage of the tridiagonal structure of $\mathbf{T}_m$). Try it on the matrix `a = gallery('wathen',20,20)` and compare your computed eigenvalues at $m = 100$ iterations (using $\mathbf{b}$ = the vector of all ones) with the true values, computed by `eig`.

Discuss the results.

Answer: See sample program. The answer should note which eigenvalues of $\mathbf{A}$ are well approximated and how accurate the approximations are.

2a. (10) Suppose we apply the (symmetric) Lanczos (tridiagonalization) algorithm (p.186) to the matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{0} \end{bmatrix}$$

where $\mathbf{C}$ is $m \times n$, $m \geq n$. Show that we obtain vectors $\mathbf{z}_1, \ldots, \mathbf{z}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_k$ satisfying

$$\begin{aligned} \mathbf{C}\mathbf{Z}_k &= \mathbf{W}_{k+1}\bar{\mathbf{T}}_{k+1}, \\ \mathbf{C}^T\mathbf{W}_k &= \mathbf{Z}_{k+1}\bar{\mathbf{T}}_{k+1} \end{aligned}$$

where $\bar{\mathbf{T}}_{k+1}$ is $(k+1) \times k$ and tridiagonal. (Hint: Write out what $\mathbf{AV} = \mathbf{VT}$ means for this particular matrix.) Now show that the eigenvalues of $\mathbf{T}_{m+n}$ are equal to the singular values of $\mathbf{C}$, the negatives of the singular values of $\mathbf{C}$, and (possibly) zeros.

Hint: Every matrix has a singular value decomposition (SVD)

$$\mathbf{C} = \mathbf{U\Sigma V}^*$$

where

- $\mathbf{U}$ has dimension $m \times m$ and $\mathbf{U}^T\mathbf{U} = \mathbf{I}$,
- $\mathbf{\Sigma}$ has dimension $m \times n$, the only nonzeros are on the main diagonal, and these *singular values* are nonnegative real numbers $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$,
- $\mathbf{V}$ has dimension $n \times n$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$.
- The columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{CC}^T$.
- The columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{C}^T\mathbf{C}$.
- The eigenvalues of $\mathbf{C}^T\mathbf{C}$ (and the nonzero eigenvalues of $\mathbf{CC}^T$) are $\sigma_1^2, \ldots, \sigma_n^2$.

Answer: It is not hard to see the first part. For the second, since $\mathbf{CV} = \mathbf{U\Sigma}$, writing the ith column of this relationship yields $\mathbf{C}\mathbf{v}_i = \sigma_i\mathbf{u}_i$. Similarly, $\mathbf{C}^T\mathbf{U} = \mathbf{\Sigma}^T\mathbf{V}$, so $\mathbf{C}^T\mathbf{u}_i = \sigma_i\mathbf{v}_i$. Now,

$$\begin{aligned} \mathbf{A}\begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} &= \begin{bmatrix} \mathbf{0} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}\mathbf{v}_i \\ \mathbf{C}^T\mathbf{u}_i \end{bmatrix} \\ &= \begin{bmatrix} \sigma\mathbf{u}_i \\ \sigma\mathbf{v}_i \end{bmatrix} \\ &= \sigma\begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix}, \end{aligned}$$

so σ_i is an eigenvalue of $\mathbf{A}$. Similarly,

$$\begin{aligned}\mathbf{A}\begin{bmatrix}\mathbf{u}_i\\ -\mathbf{v}_i\end{bmatrix} &= \begin{bmatrix}\mathbf{0} & \mathbf{C}\\ \mathbf{C}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{u}_i\\ -\mathbf{v}_i\end{bmatrix}\\ &= \begin{bmatrix}-\mathbf{C}\mathbf{v}_i\\ \mathbf{C}^T\mathbf{u}_i\end{bmatrix}\\ &= \begin{bmatrix}-\sigma\mathbf{u}_i\\ \sigma\mathbf{v}_i\end{bmatrix}\\ &= -\sigma\begin{bmatrix}\mathbf{u}_i\\ -\mathbf{v}_i\end{bmatrix},\end{aligned}$$

so $-\sigma_i$ is an eigenvalue of $\mathbf{A}$. The rank of $\mathbf{A}$ is twice the rank of $\mathbf{C}$, so all other eigenvalues of $\mathbf{A}$ are zero. Since $\mathbf{T}_{m+n}$ is similar to $\mathbf{A}$, its eigenvalues are also $\pm\sigma_i$ and zero.

2b. (10) Suppose we compute, column by column, the relations $\mathbf{CZ} = \mathbf{WB}$ and $\mathbf{C}^T\mathbf{W} = \mathbf{ZB}^T$, where $\mathbf{W}^T\mathbf{W} = \mathbf{I}$, $\mathbf{Z}^T\mathbf{Z} = \mathbf{I}$, and $\mathbf{B}$ is zero except in its main diagonal and superdiagonal. (You need not show that this $\mathbf{W}$ and $\mathbf{Z}$ are the same as those from 2a.) Show that the singular values of $\mathbf{C}$ are equal to the singular values of $\mathbf{B}_n$. Write down the recurrences for an algorithm for approximating the singular values and singular vectors of $\mathbf{C}$ by computing the first k columns of the relations and applying an SVD algorithm to $\mathbf{B}_k$.

Answer: If $\mathbf{B} = \mathbf{U\Sigma V}^T$, then

$$\mathbf{C} = \mathbf{WBZ}^T = \mathbf{WU\Sigma V}^T\mathbf{Z}^T = (\mathbf{WU})\mathbf{\Sigma}(\mathbf{ZV})^T.$$

This is the product of an orthogonal times a diagonal times an orthogonal, so this must be the SVD of $\mathbf{C}$.

The algorithm is called Lanczos bidiagonalization, and can also be derived by taking a starting vector in 2a that is zero in its second block. Let the main diagonal elements of $\mathbf{B}$ be α_i and the off-diagonal elements β_i. The recursions are

$$\mathbf{C}\mathbf{z}_i = \beta_i\mathbf{w}_{i-1} + \alpha_i\mathbf{w}_i,$$
$$\mathbf{C}^T\mathbf{w}_i = \alpha_i\mathbf{z}_i - \beta_{i+1}\mathbf{z}_{i+1}.$$

Given $\mathbf{z}_1$ of norm one, use the first relation to compute $\mathbf{w}_1$, with α_i $(i = 1)$ chosen to make $\mathbf{w}_i^T\mathbf{w}_i = 1$. Use the second relation to compute $\mathbf{z}_2$, choosing β_{i+1} to enforce $\mathbf{z}_{i+1}^T\mathbf{z}_{i+1} = 1$. Repeat for $i = 2, \ldots$. Show by induction that the bases are orthonormal.

2c. (10) Implement your algorithm from 2b in Matlab and try it on the matrix from `load('west0479.mat')`. Compare the singular values computed for $k = 100$ to the true values computed using `svd(full(west0479))`. Discuss the results.

Answer: See sample code.