December 21, 2011

Exercises, day 13

Exercise 13.1. **Arnoldi relations.**

Let $\mathbf{A}$ be an $n\times n$ matrix. Assume that

$$\mathbf{A}\mathbf{W}_\ell=[\mathbf{W}_\ell,\mathbf{v}_{\ell+1}]\,\underline{G}_\ell \tag{1}$$

for some $n\times\ell$ matrix $\mathbf{W}_\ell$ and n-vector $\mathbf{v}_{\ell+1}$ for which $[\mathbf{W}_\ell,\mathbf{v}_{\ell+1}]$ is orthonormal and an $(\ell+1)\times\ell$ matrix $\underline{G}_\ell$. Note that $\underline{G}$ is not assumed to be Hessenberg.

(a) Show that there is an $\ell\times\ell$ unitary matrix U, $U=Q_1\cdot\ldots\cdot Q_{\ell-1}$ a product of $\ell-1$ Householder reflections (i.e., $Q_j=1-2q_jq_j^*$ for some normalized ℓ-vector q_j) such that

$$\underline{H}_\ell\equiv\left[\begin{array}{cc}U^* & \vec{0}\\ \vec{0}^* & 1\end{array}\right]\underline{G}_\ell\,U$$

is $(\ell+1)\times\ell$ upper Hessenberg.

(b) Conclude that, with $\mathbf{V}_\ell\equiv\mathbf{W}_\ell U$, we have the Arnoldi relation

$$\mathbf{A}\mathbf{V}_\ell=[\mathbf{V}_\ell,\mathbf{v}_{\ell+1}]\,\underline{H}_\ell.$$

Note that the vector $\mathbf{v}_{\ell+1}$ did not "change". Note that the columns of $\mathbf{V}_\ell$ are a 'rotated' but orthonomale basis for the space spanned by the columns of $\mathbf{W}_\ell$.

(c) Note that $\mathbf{V}_\ell=\mathbf{W}_\ell\,(Q_1\ldots Q_{\ell-1})$ can also be obtained as $\mathbf{V}_\ell=(\ldots(\mathbf{W}_\ell Q_1)Q_2)\ldots Q_{\ell-1})$. Which way of computing $\mathbf{V}_\ell$ do you prefer?

(d) Let $\mathcal{V}_\ell$ be an ℓ dimensional subspace such that $\mathbf{A}\mathcal{V}_\ell\equiv\{\mathbf{A}\mathbf{v}\,\big|\,\mathbf{v}\in\mathcal{V}_\ell\}\subset\mathcal{V}_\ell+[\mathbf{w}_{\ell+1}]$ for some vector $\mathbf{w}_{\ell+1}$. Here, $\mathcal{V}_\ell+[\mathbf{w}_{\ell+1}]$ is the space spanned by the vector $\mathbf{w}_{\ell+1}$ and vectors from $\mathcal{V}_\ell$. Show that $\mathcal{V}_\ell$ is a Krylov subspace generated by $\mathbf{A}$ and some vector $\mathbf{v}_1$. Describe a procedure to compute $\mathbf{v}_1$.

Exercise 13.2. **Implicitly restarted Arnoldi method.**

Let $\mathbf{A}$ be an $n\times n$ matrix. Assume that

$$\mathbf{A}\mathbf{V}_k=\mathbf{V}_{k+1}\underline{H}_k$$

is a partial Arnoldi decomposition: $\mathbf{V}_j$ is $n\times j$ orthonormal, $\underline{H}_k$ is $(k+1)\times k$ upper Hessenberg. Let H_k be the $k\times k$ upper block of $\underline{H}_k$. Let

$$H_k=Q_kS_kQ_k^*$$

be the Schur decomposition of H_k, that is, S_k is $k\times k$ upper triangular and Q_k is $k\times k$ unitary (see Exercise 3.15). Let $\vartheta_1,\ldots,\vartheta_\ell,\vartheta_{\ell+1},\ldots,\vartheta_k$ be the on the diagonal of S_k (see Exercise 13.5 below).

(a) Prove that the ϑ_i are Ritz values with pre-Ritz vectors $y_i\equiv Q_ke_i$ and Ritz vector $\mathbf{u}_i\equiv\mathbf{V}_ky_i$.

Assume we want to continue Arnoldi's process with the space $\mathcal{V}_\ell$ spanned by the ℓ Ritz vectors associated to the first ℓ Ritz values $\vartheta_1,\ldots,\vartheta_\ell$.

(b) Let $\mathbf{W}_\ell$ be the matrix consisting of the first ℓ columns of $\mathbf{V}_kQ_k$. Show that the columns of $\mathbf{W}_\ell$ form an orthonormal basis if $\mathcal{V}_\ell$.

(c) Show that $\mathbf{A}\mathbf{W}_\ell=[\mathbf{W}_\ell,\mathbf{v}_{k+1}]\underline{G}$ for some $(\ell+1)\times\ell$ matrix $\underline{G}$. Show that the $\ell\times\ell$ upper block of $\underline{G}$ is equal to the $\ell\times\ell$ left upper block of S.

(d) Use Exercise 13.1 to turn $\mathbf{A}\mathbf{W}_\ell=[\mathbf{W}_\ell,\mathbf{v}_{k+1}]\,\underline{G}$ into a partial Arnoldi relation

$$\mathbf{A}\mathbf{V}_\ell'=[\mathbf{V}_\ell',\mathbf{v}_{\ell+1}']\,\underline{H}_\ell{}',$$

i.e., the columns of $\mathbf{V}'_\ell$ form an orthonormal Krylov basis for $\mathcal{V}_\ell$. What can you say of $\mathbf{v}'_{\ell+1}$?

Exercise 13.3. **Arnoldi and polynomial filters.**

Let $\mathbf{A}$ be an $n \times n$ matrix. Assume that

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k, \quad \mathbf{v}_1 \equiv \mathbf{V}_k e_1,$$

is a partial Arnoldi decomposition: $\mathbf{V}_j$ is $n \times j$ orthonormal, $\underline{H}_k$ is $(k+1)\times k$ upper Hessenberg. Let H_k be the $k \times k$ upper block of $\underline{H}_k$.

In this exercise, we will show that applying a few steps of the shifted QR-algorithm to the matrix $\underline{H}_k$ has the same effect as applying a polynomial filter in $\mathbf{A}$ and $\mathbf{v}_1$, that is, taken $P(\mathbf{A})\mathbf{v}_1$ as an initial vector for the Arnoldi process, rather than $\mathbf{v}_1$.

Select a shift $\mu \in \mathbb{C}$. Form the QR-decomposition of $\underline{H}_k - \mu\,\underline{I}_k$:

$$\underline{H}_k - \mu\,\underline{I}_k = \underline{Q}_k R_k$$

where $\underline{Q}_k$ is $(k+1)\times k$ orthonormal and R_k is $k \times k$ upper triangular.

(a) Show that such a decomposition exists. Show that $\underline{Q}_k$ is upper Hessenberg.

Let $\underline{Q}_{k-1}$ be the $k \times (k-1)$ left upper block of $\underline{Q}_k$. Define

$$\underline{H}^+_{k-1} \equiv R_k \underline{Q}_{k-1} + \mu\,\underline{I}_{k-1} \quad \text{and} \quad \mathbf{V}^+_{k-1} \equiv \mathbf{V}_k \underline{Q}_{k-1}, \quad \mathbf{V}^+_k \equiv \mathbf{V}_{k+1}\underline{Q}_k.$$

(b) Prove that $\underline{H}^+_{k-1}$ is $k \times (k-1)$ upper Hessenberg. Prove that

$$\underline{Q}_k \underline{H}^+_{k-1} = \underline{H}_k \underline{Q}_{k-1} \quad \text{and} \quad \mathbf{A}\mathbf{V}^+_{k-1} = \mathbf{V}^+_k \underline{H}^+_{k-1}, \quad (\mathbf{A} - \mu\,\mathbf{I})\mathbf{V}_k = \mathbf{V}^+_k R_k.$$

(c) Show that

$$(\mathbf{A} - \mu\,\mathbf{I})\mathbf{v}_1 = \tau\,\mathbf{v}^+_1$$

for some scalar τ. Here, $\mathbf{v}_1$ and $\mathbf{v}^+_1$ are the first columns of the matrix $\mathbf{V}_k$ and $\mathbf{V}^+_k$, respectively. Conclude that the columns of $\mathbf{V}^+_k$ form an orthonormal Krylov basis for the Krylov subspace of order k generated by $\mathbf{A}$ and $(\mathbf{A} - \mu\,\mathbf{I})\mathbf{v}_1$.

(d) Let $P(\lambda) \equiv (\lambda - \mu_1)\cdot \ldots \cdot (\lambda - \mu_{k-\ell})$ a polynomial. Apply the above procedure repeatedly to find an partial Arnoldi decomposition

$$\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\,\underline{H}'_\ell, \qquad \widetilde{\mathbf{v}}_1 \equiv \widetilde{\mathbf{V}}_\ell e_1 = \widetilde{\tau}\,P(\mathbf{A})\mathbf{v}_1$$

such that the columns of $\widetilde{\mathbf{V}}_{\ell+1}$ form an orthonormal Krylov basis of the Krylov subspace generated by $\mathbf{A}$ and $\widetilde{\mathbf{v}}_1$. $\widetilde{\tau}$ is some scalar.

(e) Explain how you can compute $\widetilde{\mathbf{V}}_{\ell+1}$ by applying only one $(k+1)\times(\ell+1)$ orthonormal matrix $\widetilde{Q}$ from the right to $\mathbf{V}_{k+1}$: $\widetilde{\mathbf{V}}_{\ell+1} = \mathbf{V}_k \widetilde{Q}$.

Exercise 13.4. **Implicitly restarted Arnoldi method 2.**

We continue Exercise 13.3. We will see that the *implicit* polynomial filtering approach from Exercise 13.3 will provide us with an elegant (and implicit) way for restarting Arnoldi (see Exercise 13.2) with the space of selected Ritz vectors.

Let $\vartheta_1, \ldots, \vartheta_k$ be the eigenvalues of H_k. Apply the steps in (a) and (b) of Exercise 13.3 with $\mu = \vartheta_k$.

(a) Show that $\underline{Q}_k e_k = e_{k+1}$. (Hint: note that $\mathbf{V}_k e_k = \mathbf{V}_{k+1} e_{k+1}$).

(b) Show that the eigenvalues of the $(k-1)\times(k-1)$ upper block H^+_{k-1} of $\underline{H}^+_{k-1}$ are equal to $\vartheta_1, \ldots, \vartheta_{k-1}$.

(c) Now, take $P(\lambda) = (\lambda - \vartheta_{\ell+1})\ldots(\lambda - \vartheta_k)$. Show that the columns of $\widetilde{\mathbf{V}}_\ell$ form an orthonormal basis for the space spanned by the Ritz vectors associated to the Ritz values $\vartheta_1, \ldots, \vartheta_\ell$.

Exercise 13.5. **Sorting the Schur decomposition.**

Consider the (complex) 2×2 matrix $\begin{bmatrix} \lambda & \alpha \\ 0 & \mu \end{bmatrix}$.

(a) Show there is a Givens rotation $\begin{bmatrix} c & s \\ -\bar{s} & c \end{bmatrix}$ with $c\in[0,1]$ en $s\in\mathbb{C}$ such that $c^2+|s|^2=1$ ($c=\cos(\phi)$ and $s=\sin(\phi)$) such that

$$\begin{bmatrix} c & s \\ -\bar{s} & c \end{bmatrix}\begin{bmatrix} \lambda & \alpha \\ 0 & \mu \end{bmatrix}\begin{bmatrix} c & -s \\ \bar{s} & c \end{bmatrix}=\begin{bmatrix} \mu & \beta \\ 0 & \lambda \end{bmatrix}$$

for some appropriate scalar β. With $t\equiv s/c$ we should have that $t=(\mu-\lambda)/\alpha$. Hence $c=1/\sqrt{|t|^2+1}$ and $s=ct$.

(b) Let S be an $k\times k$ upper triangular matrix. Suppose we want to 'switch' the 1th and the kth diagonal element: we are interested in a unitary matrix Q such that $S'\equiv Q^*SQ$ is upper triangular such that $S'_{ii}=S_{ii}$ if $i\notin\{1,k\}$ and $S'_{11}=S_{kk}$ and $S'_{kk}=S_{11}$. Show that this can be achieve with Givens rotations. How many Givens rotations are required?

Any permutation can be expressed as a product of basic permutation, where a basic permutation is a permutation that switches two indices only.

(c) Let P be a permutation of $(1,\ldots,k)$ (i.e., P is a bijection onto $\{1,2,\ldots,k\}$). Show that there is a unitary matrix Q such that with $S'\equiv Q^*SQ$, S' is upper triangular, and $\mathrm{diag}(S')=\mathrm{diag}(S)(P)$. Q can be obtained as a product of Givens rotations.

(d) Write a Matlab function subroutine `[S1,Q1]=SortSchur(S,Q,p);` that sorts the Schur decomposition $H=QSQ^*$ according the permutation P: the matrices are $k\times k$, Q is unitary, S is upper triangular such that $H=QSQ^*$ (ex. `[Q,S]=schur(H);`), then the matrices $S_1\equiv$`S1` and $Q_1\equiv$`Q1` are such that S_1 is upper triangular, Q_1 is unitary, $H=Q_1S_1Q_1^*$ and with `D=diag(S); D1=diag(S1);` we have that `D1=D(P);`.

Exercise 13.6. **Deflation.**

Exercise 13.7. **Rational Krylov space.**

Exercise 13.8. **Rational Krylov sequence method.**

Exercise 13.9. **Jacobi-Davidson.**