TEKNISKA HÖGSKOLAN I LINKÖPING
Matematiska institutionen
Beräkningsmatematik/Fredrik Berntsson

Exam TANA15 Numerical Linear Algebra, Y4, Mat4

Datum: 20:e Mars, 2024.

Hjälpmedel:

1. Föreläsningsanteckningar utskrivna från kurshemsidan utan egna anteckningar.
2. Räknedosa i fickformat, med nollställt minne och utan instruktionsbok.

Examinator: Fredrik Berntsson

Maximalt antal poäng: 25 poäng. För godkänt krävs 10 poäng.

Jourhavandelärare Fredrik Berntsson (telefon 013 28 28 60)

Good luck!

(4p) **1:** Do the following

a) Prove that $\dim(\mathrm{Null}(A)) + \dim(\mathrm{Range}(A)) = n$.
Hint Pick a basis for $\mathrm{Null}(A)$ and complete to a basis for all of $\mathbb{R}^n$.

b) Prove the inequality $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$.

c) Show that if P is an orthogonal projection and λ is an eigenvalue of P then λ is either 0 or 1.

(4p) **2:** Let,

$$f(x) = \begin{pmatrix} \exp(2x_1 + x_2) - 2 \\ \sin(3x_1 - x_2) \end{pmatrix}.$$

a) Compute the Jacobian J_f and formulate the Newton method for finding a root of the equation $f(x) = 0$.

b) Let $x^{(0)} = (0\,,\,0)^T$ and perform one step of the Newton metod and compute the next iterate $x^{(1)}$.

(4p) **3:** Suppose A is an $m \times n$, $m > n$, matrix and the linear system $Ax = b$ doesn't have a exact solution. The *Total least squares* solution x satisfies $(A + E)x = b + r$, where $[E, r]$ is given by

$$\min \|[E, r]\|_2 \text{ such that } (A + E)x = b + r.$$

Do the following:

a) Show that the *Total least squares* problem always has a solution.

b) Use the singular value decomposition to derive the solution to the problem. Note that it may not always be possible to find the Total least squares solution using the singular value decomposition and in the case it fails you should give a clear criteria that shows if the formula worked or not.

(4p) **4:** Any matrix $A \in \mathbb{R}^{m\times n}$, $m > n$, has a *singular value decomposition* $A = U\Sigma V^T$. Do the following:

a) Consider a linear system $Ax = b$, $m > n$, where $\text{rank}(A) = k < n$. Use the SVD to a basis for the both the range $\text{Range}(A)$ and its orthogonal complement. Also give a criteria that guarantees that a solution to the linear system exists. Your criteria should be expressen in terms of the basis vectors and the vector b. Also the criteria should be efficient to check for the case when $k \approx n \approx m$.

b) Consider the linear system $A^T x = b$, where as before $\text{rank}(A) = k < n$. Provide a criteria for existance of a solution to the linear system expressed in terms of b and the singular vectors. Also write down the formula for the solution x. Is the solution unique? Motivate clearly.

(4p) **5:** Do the following:

a) Clearly demonstrate how a bidiagonal reduction $A = UBV^T$ can be computed using Householder reflections. You have to specify which elements of the matrix are used to create each reflection. It is enough to consider the 4×4 case.

b) Give the definition of the singular values of an $m \times n$, $m > n$, matrix A. Also suppose we have all the eigenvalues $\{\lambda_i\}$ of $B^T B$, where $A = UBV^T$ is the bidiagonal reduction. Clearly show how to obtain the singular values of A in terms of the eigenvalues of B. What are the dimensions of the matrices B and $B^T B$?

(5p) **6:** **a)** Show that any matrix $A \in \mathbb{R}^{n\times n}$ can be factorized as $A = QTQ^H$, where Q is unitary and T upper triangular. This is called the *Schur decomposition.*

b) A matrix B is called non-defective if it has a full set of eigenvectors, i.e. the decomposition $B = XDX^{-1}$ exists. Use the Shur decomposition to prove that if A is defective then for any $\varepsilon > 0$ there is a non-defective matrix B such that $\|A - B\|_2 \leq \varepsilon$.

Remark From **b)** we conclude that if a matrix is supposed to be defective and we compute a numerical approximation it is likely that the matrix turns out to be non-defective due to round-off errors.

1: For **a)** suppose the dimension of the null space is k. Then there is a basis $\{x_1, x_2, \ldots, x_k\}$ for the nullspace. Add $n-k$ linearly independent vectors $\{\widetilde{x}_{k+1}, \ldots, \widetilde{x}_n\}$ so that we have a basis for $\mathbb{R}^n$. Now take a vector y that belongs to the subspace Range(A), i.e. $y = Ax$ for some $x \in \mathbb{R}^n$. We can express x using the above basis and since $Ax_i = 0$, for $i = 1, \ldots, k$, we find that y is a linear combination of the vectors $\{A\widetilde{x}_{k+1}, \ldots, A\widetilde{x}\}$. So the dimension is at most $n-k$. To show that the dimension is exactly $n-k$ we assume that there is a linear combination so that

$$0 = \sum_{i=k+1}^{n} c_i A\widetilde{x}_i = A(\sum_{i=k+1}^{n} c_i A\widetilde{x}_i) = Az,$$

so z belongs to the nullspace which contradicts the assumption that the set of vectors

$$\{x_1, x_2, \ldots, x_k, \widetilde{x}_{k+1}, \ldots, \widetilde{x}_n\}$$

was a basis. So the dimension of the range is exactly $n-k$.

For **b)** we demonstrate the first inequality by

$$\|x\|_\infty^2 = \max_{1\le i\le n} |x_i|^2 \le \sum_{i=1}^{n} |x_i|^2 = \|x\|_2^2.$$

Also, since $|x_i| \le \|x\|_\infty$, we have

$$\|x\|_2^2 = \sum_{i=1}^{n} |x_i|^2 \le \sum_{i=1}^{n} \|x\|_\infty^2 = n\|x\|_\infty.$$

For **c)** we simply note that λ is an eigenvalue of P is there is a vector $x \neq 0$ such that $Px = \lambda x$. The property that signify projections is that $Px = x$ for all $x \in$ Range(P). Thus for such x we get $x = \lambda x$ showing that 1 is an eigenvalue. For all vectors x that does not belong to Range(P) the equation is $Px = \lambda x$. Since Px obviously belongs to the range this is only possible if $\lambda = 0$.

2: For **a)** we recall that $(J_f)_{ij}(x) = (\partial_{x_j} f_i(x))$. Thus

$$J_f(x) = \begin{pmatrix} 2\mathrm{e}^{3x_1+x_2} & \mathrm{e}^{3x_1+x_2} \\ 3\cos(3x_1 - x_2) & -\cos(3x_1 - x_2) \end{pmatrix},$$

where $x = (x_1, x_2)^T$. The Newton method is formulated as follows: Solve $J_f(x^{(n)})s^{(n)} = -f(x^{(n)})$,and update $x^{(n+1)} = x^{(n)} + x^{(n)}$.

For **b)** we evaluate $f(x^{(0)}) = f((0,0)^T) = (1,0)^T$, and

$$J_f((0,0)^T) = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}.$$

In the Newton step we first solve the linear system $J_f s^{(0)} = -f(x^{(0)}$, or

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} s^{(0)} = \begin{pmatrix} 3 \\ -1 \end{pmatrix},$$

which gives $s^{(0} = (0.2, 0.6)$. Thus $x^{(1)} = x^{(0)} + s^{(0)} = (0.2, 0.6)^T$.

3: For **a)** we simply observe that the equation $(A+E)x = b+r$ is satisfied, for any x, if $E = -A$ and $r = -b$. The minimum is also bounded from below (by 0). Thus there is some E, r that gives the minimum.

For **b)** we can assume that the agumented matrix $[A, b]$ has full rank since otherwise the minimum would be zero and the linear system $Ax = b$ have a solution. We then compute the singular value decomposition $[A, b] = U\Sigma V^T$ of the $m \times (n+1)$ matrix. The smallest perturbation $[E, r]$ that makes the matrix $[A+E, b+r]$ rank deficient is given by the last singular component $[E, r] = -\sigma_{n+1} u_{n+1} v_{n+1}^T$. There is an x such that $(A+E)x = (b+r)$ if $[A+E, b+r](x, -1)^T = 0$, i.e. $(x, -1)^T$ belongs to the null space of $[A+E, b+r]$. By the construction above the null space is exactly v_{n+1}. So we just take the last singular vector and multiply by a constant so that the last component becomes 1. Thus $x = v_{n+1}(1:n)/v_{n+1}(n+1)$. This is the total least squares solution.

This obviously fails if $v_{n+1}(n+1) = 0$. In that case we have to figure out something else to find the total least squares solution.

4: For **a)** we remark that we can write A in the form

$$A = \sum_{i=1}^{k} \sigma_i u_i v_i^T.$$

Here we clearly see that $\text{Range}(A) = \text{span}(u_1, \ldots, u_k)$. The orthogonal complement is $\text{Range}(A)^\perp = \text{span}(u_{k+1}, \ldots, u_m)$. Existance of solution means that $b \in \text{Range}(A)$ which means b doesn't have a component in $\text{Range}(A)^\perp$. For large k the easiest way to check this is $u_i^T b = 0$, for $i = k+1, \ldots, m$.

For **b)** we simply apply the transpose to the above formula for A to obtain

$$A^T = \sum_{i=1}^{k} \sigma_i v_i u_i^T.$$

This means that now we have $\text{Range}(A^T) = \text{span}(v_1, \ldots, v_k)$. A criteria for existance is thus $v_i^T b = 0$, for $i = k+1, \ldots, n$. If this criteria is satisfied we can write

$$b = \sum_{i=1}^{k} (v_i^T b) v_i = A^T x = \sum_{i=1}^{k} \sigma_i (u_i^T x) v_i.$$

Identifying coefficients gives us $v_i^T b = \sigma_i(u_i^T x)$, for $i = 1, \dots, k$. We can express x in the basis $\{u_1, \dots, u_m\}$ so

$$x = \sum_{i=1}^{m} (u_i^T x) u_i = \sum_{i=1}^{k} \frac{v_i^T b}{\sigma_i} u_i + \sum_{i=k+1}^{m} c_i u_i$$

where c_i are free parameters. The solution is not unique.

5: For **a)** we illustrate the algorithm as follows: First we use a reflection H_1 applied from the left. The reflection is selected so the elements $A(2:4,1)$ are set to zero. Second we apply a reflection H_2 from the right to zero out the elements $\widetilde{A}(1,3:4)$. We get

$$H_1 \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \cdot \begin{pmatrix} + & + & + & + \\ 0 & + & + & + \\ 0 & + & + & + \\ 0 & + & + & + \end{pmatrix} H_2^T = \begin{pmatrix} x & + & 0 & 0 \\ 0 & + & + & + \\ 0 & + & + & + \\ 0 & + & + & + \end{pmatrix}.$$

Now we continue with reflections H_3 and H_4 that zero out $A(3:4,2)$ and $A(2,4)$. We get

$$H_3 \begin{pmatrix} x & x & 0 & 0 \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix} H_4^T = \begin{pmatrix} x & x & 0 & 0 \\ 0 & + & + & + \\ 0 & 0 & + & + \\ 0 & 0 & + & + \end{pmatrix} H_4^T = \begin{pmatrix} x & x & 0 & 0 \\ 0 & x & + & 0 \\ 0 & 0 & + & + \\ 0 & 0 & + & + \end{pmatrix}.$$

Finally we apply one reflection H_5 from the left to zero out the element $A(4,3)$. We get

$$H_5 \begin{pmatrix} x & x & 0 & 0 \\ 0 & x & x & 0 \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{pmatrix} = \begin{pmatrix} x & x & 0 & 0 \\ 0 & x & x & 0 \\ 0 & 0 & + & + \\ 0 & 0 & 0 & + \end{pmatrix},$$

which is bidiagonal.

For **b)** there easiest way to define the singular values is to say that the singular value decomposition is $A = U\Sigma V^T$, where U and V are orthogonal matrices and Σ is diagonal. The singular values σ_k are the diagonal elements of Σ provided that U and V are chosen so that the diagonal elements are positive and sorted in descending order. The dimension of $B^T B$ is $n \times n$ and the dimension of BB^T is $m \times m$. If $A = UBV^T$ then $A^T A = UB^T BU^T$ so the eigenvalues of $B^T B$ are the same as those of $A^T A$. Also suppose $A = \bar{U}\Sigma\bar{V}^T$ is the singular value decomposition of A. Then $A^T A = \bar{V}\Sigma^T\Sigma\bar{V}^T$. So the eigenvalues of $A^T A$ are $\lambda_i = \sigma_i^2$, where σ_i are the singular values of A. Thus $\sigma_i = \sqrt{\lambda_i}$, $i = 1, 2, \dots, n$. We are just missing $m - n$ zero singular values to get the correct dimension.

6: For **a)** we pick an eigenpair (λ, x). If we compute the full QR decomposition of $x \in \mathbb{R}^{n\times 1}$ we obtain an orthogonal matrix suxch that $Q = (x, Q_2)$, where $Q_2^H x = 0$. This is assuming that $\|x_1\|_2 = 1$. We find that

$$Q^H AQ = (x, Q_2)^T A(x, Q_2) = (x, Q_2)^H (Ax, AQ_2) = (x, Q_2)^H (\lambda x, AQ_2) =$$

$$\begin{pmatrix} \lambda x^H x & x^H A Q_2 \\ \lambda Q_2^H x & Q_2^H A Q_2 \end{pmatrix} = \begin{pmatrix} \lambda & w^H \\ 0 & B \end{pmatrix},$$

where we have the correct structure. This is the first step of finding the Hessenberg decomposition. Now we make the induction argument that the Hessenberg decomposition exists for dimension $n-1$ and find $B = Q_1 H_1 Q_1^H$. We then have

$$Q^H A Q = \begin{pmatrix} \lambda & w^H \\ 0 & Q_1 H_1 Q_1^H \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} = \begin{pmatrix} \lambda & w^H \\ 0 & H_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}^H.$$

For **b)** we simply note that $A^H = (QTQ^H)^H = QT^HQ^H$. For symmetric matrices, i.e. A real and $A^T = A$, we thus get $A^T = A^H = QT^HQ^H = A = QTQ^H$. Thus $T^H = T$ which means that T is a diagonal since we already knew that T is upper triangular. Also the diagonal elements satisfy $(T)_{ii} = (\bar{T})_{ii}$ which means the elements on the diagonal are real. Since the diagonal elements of T are also the eigenvalues of A this shows that the eigenvalues are real.

For **c)** we assume that A is defective and compute its Shur decomposition $A = QTQ^H$. For A to be defective it has to have at least one eigenvalue λ_1 with an algebraic multiplicity $\gamma_1(\lambda_1)$ strictly larger than the geometric multiplicity $\gamma_2(\lambda_1)$. Thus, if all diagonal elements of T were different then the matrix A would be non-defective. Thus we pick a diagonal matrix $D = \text{diag}(\epsilon_1, \ldots, \epsilon_n)$ so that $T + D$ has unique diagonal elements. Then $B = Q(T+D)Q^H$ is non-defective and $\|A-B\|_2 = \|D\|_2 \le \max|\epsilon_i| = \epsilon$.