UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Examination in: MAT-INF 4130 — Numerical Linear Algebra.

Day of examination: Monday 14. December 2015.

Examination hours: 9:00 – 13:00.

This problem set consists of 5 pages.

Appendices: None.

Permitted aids: None.

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

All 10 part questions will be weighted equally.

Problem 1.

a) Let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 3 \end{bmatrix}$$

Compute $\|\mathbf{A}\|_1$ and $\|\mathbf{A}\|_\infty$.

Solution: $\|\mathbf{A}\|_1$ is the maximum absolute column sum in $\mathbf{A}$. Since the absolute column sums are 2 and 6, we have that $\|\mathbf{A}\|_1 = 6$.

$\|\mathbf{A}\|_\infty$ is the maximum absolute row sum in $\mathbf{A}$. Since the absolute row sums are 3, 1, and 4, we have that $\|\mathbf{A}\|_\infty = 4$.

b) Let $\mathbf{B}$ be the matrix

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Find the spaces $\text{span}(\mathbf{B}^T)$ and $\ker(\mathbf{B})$.

Solution: we have that

$$\mathbf{B}^T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

Clearly the two columns here are linearly independent, so that they are a base for $\text{span}(\mathbf{B}^T)$. To find $\ker(\mathbf{B})$ first perform row reduction:

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}.$$

x_3 is thus a free variable, and we must have that $x_1 = x_3$, and $x_2 = -2x_3$

(Continued on page 2.)

for any $\mathbf{x} \in \ker \mathbf{B}$. Therefore

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

so that $(1, -2, 1)$ is a basis for $\ker \mathbf{B}$.

c) Consider the underdetermined linear system

$$\begin{array}{llll} x_1 & & -x_3 & = 4 \\ x_1 & +x_2 & +x_3 & = 12 \end{array}$$

Find the solution $\mathbf{x} \in \mathbb{R}^3$ with $\|\mathbf{x}\|_2$ as small as possible.

Solution: There are several ways one can solve this task.

We can use that $\text{span}(\mathbf{B}^T)$ and $\ker(\mathbf{B})$ is an orthogonal basis for $\mathbb{R}^3$. Any solution $\mathbf{x}$ to the above equation can thus be written as $\mathbf{x} = \mathbf{B}^T\mathbf{y} + \mathbf{z}$, where $\mathbf{y} \in \mathbb{R}^2$, $\mathbf{z} \in \ker \mathbf{B}$. The solution with minimum Euclidean norm is then $\mathbf{x} = \mathbf{B}^T\mathbf{y}$, since $\text{span}(\mathbf{B}^T)$ and $\ker(\mathbf{B})$ are orthogonal, so that one can solve for $\mathbf{y}$ first in $\mathbf{B}\mathbf{B}^T\mathbf{y} = \begin{bmatrix} 4 \\ 12 \end{bmatrix}$. Since $\mathbf{B}\mathbf{B}^T = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, we get that $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, and finally $\mathbf{x} = \mathbf{B}^T\mathbf{y} = (6, 4, 2)$.

We can also go as follows: We have that

$$\begin{bmatrix} 1 & 0 & -1 & 4 \\ 1 & 1 & 1 & 12 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 4 \\ 0 & 1 & 2 & 8 \end{bmatrix},$$

so that the general solution is

$$\mathbf{x} = \begin{bmatrix} x_3 + 4 \\ -2x_3 + 8 \\ x_3 \end{bmatrix}.$$

We have that $\|\mathbf{x}\|_2^2 = (x_3+4)^2 + (-2x_3+8)^2 + x_3^2 = 6x_3^2 - 24x_2 + 80$. This is minimized when $12x_3 - 24 = 0$, i.e. when $x_3 = 2$, which gives $\mathbf{x} = (6, 4, 2)$.

It is also rather straightforward to solve this exercise using pseudoinverses. Consider $\mathbf{B}\mathbf{B}^T = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ (rather than $\mathbf{B}^T\mathbf{B}$, which is a 3×3 matrix). We see that the singular values of $\mathbf{B}$ are $\sigma_1 = \sqrt{3}$ and $\sigma_2 = \sqrt{2}$. The corresponding eigenvectors for $\mathbf{B}\mathbf{B}^T$ are $\mathbf{e}_2$ and $\mathbf{e}_1$, respectively. Since $\frac{1}{\sigma_1}\mathbf{B}^T\mathbf{e}_1 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ and $\frac{1}{\sigma_2}\mathbf{B}^T\mathbf{e}_2 = (1/\sqrt{2}, 0, -1/\sqrt{2})$, a singular value factorization of $\mathbf{B}^T$ is

$$\mathbf{B}^T = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Transposing this we get a singular value factorization for $\mathbf{B}$, and we then easily get the following expression for the pseudoinverse:

$$\mathbf{B}^\dagger = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/3 \\ 0 & 1/3 \\ -1/2 & 1/3 \end{bmatrix}.$$

(Continued on page 3.)

The least squares solution (which also here is a solution) with minimum Euclidean norm can now be obtained by computing

$$\mathbf{B}^{\dagger}\mathbf{b} = \begin{bmatrix} 1/2 & 1/3 \\ 0 & 1/3 \\ -1/2 & 1/3 \end{bmatrix} \begin{bmatrix} 4 \\ 12 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}.$$

d) Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a matrix with linearly independent columns, and $\mathbf{b} \in \mathbb{R}^m$ a vector. Assume that we use the Gauss-Seidel method to solve the normal equations $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$. Will the method converge? Justify your answer.

Solution: If $\mathbf{A}$ has linearly independent columns, $\mathbf{A}^T\mathbf{A}$ is invertible (by the characterization of least squares solutions in terms of the normal equations), so that it is also positive definite. But from Theorem 11.15 we know that the Gauss-Seidel method converges for any positive definite matrix.

Problem 2.

a) Let $\mathbf{E} \in \mathbb{R}^{n\times n}$ be of the form $\mathbf{E} = \mathbf{I} + \mathbf{u}\mathbf{u}^T$, where $\mathbf{u} \in \mathbb{R}^n$. Show that $\mathbf{E}$ is symmetric and positive definite, and find an expression for $\mathbf{E}^{-1}$.
(Hint: $\mathbf{E}^{-1}$ is of the form $\mathbf{E}^{-1} = \mathbf{I} + a\mathbf{u}\mathbf{u}^T$ for some $a \in \mathbb{R}$.)

Solution: We have that $\mathbf{E}^T = (\mathbf{I} + \mathbf{u}\mathbf{u}^T)^T = \mathbf{I}^T + (\mathbf{u}\mathbf{u}^T)^T = \mathbf{I} + \mathbf{u}\mathbf{u}^T = \mathbf{E}$, and

$$\mathbf{x}^T\mathbf{E}\mathbf{x} = \mathbf{x}^T(\mathbf{I} + \mathbf{u}\mathbf{u}^T)\mathbf{x} = \mathbf{x}^T\mathbf{x} + \mathbf{x}^T\mathbf{u}\mathbf{u}^T\mathbf{x} = \|\mathbf{x}\|^2 + (\mathbf{x}^T\mathbf{u})^2 > 0,$$

so that $\mathbf{E}$ is symmetric and positive definite. Using the hint we compute

$$(\mathbf{I}+a\mathbf{u}\mathbf{u}^T)(\mathbf{I}+\mathbf{u}\mathbf{u}^T) = \mathbf{I}+(1+a)\mathbf{u}\mathbf{u}^T + a\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T = \mathbf{I}+(1+a+a\|\mathbf{u}\|^2)\mathbf{u}\mathbf{u}^T.$$

This equals $\mathbf{I}$ if $1 + a + a\|\mathbf{u}\|^2 = 0$, i.e. if $a = -1/(1 + \|\mathbf{u}\|^2)$. This shows that

$$\mathbf{E}^{-1} = \mathbf{I} - \frac{1}{1 + \|\mathbf{u}\|^2}\mathbf{u}\mathbf{u}^T.$$

b) Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be of the form $\mathbf{A} = \mathbf{B}+\mathbf{u}\mathbf{u}^T$, where $\mathbf{B} \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, and $\mathbf{u} \in \mathbb{R}^n$. Show that $\mathbf{A}$ can be decomposed on the form

$$\mathbf{A} = \mathbf{L}(\mathbf{I} + \mathbf{v}\mathbf{v}^T)\mathbf{L}^T,$$

where $\mathbf{L}$ is nonsingular and lower triangular, and $\mathbf{v} \in \mathbb{R}^n$.

Solution: Since $\mathbf{B}$ is symmetric and positive definite it has a Cholesky factorization $\mathbf{B} = \mathbf{L}\mathbf{L}^T$. We have that

$$\mathbf{L}(\mathbf{I} + \mathbf{v}\mathbf{v}^T)\mathbf{L}^T = \mathbf{L}\mathbf{L}^T + \mathbf{L}\mathbf{v}\mathbf{v}^T\mathbf{L}^T = \mathbf{B} + \mathbf{L}\mathbf{v}(\mathbf{L}\mathbf{v})^T.$$

If we now choose $\mathbf{v}$ so that $\mathbf{L}\mathbf{v} = \mathbf{u}$ (this is possible since $\mathbf{L}$ is nonsingular), this equals $\mathbf{B} + \mathbf{u}\mathbf{u}^T = \mathbf{A}$, and this shows that $\mathbf{A}$ can be written on the desired form.

c) Assume that the Cholesky decomposition of $\mathbf{B}$ is already computed. Outline a procedure to solve the system $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{A}$ is of the form above.

Solution: We first find a $\mathbf{v}$ so that $\mathbf{A} = \mathbf{L}(\mathbf{I} + \mathbf{v}\mathbf{v}^T)\mathbf{L}^T$ (by solving $\mathbf{L}\mathbf{v} = \mathbf{u}$, which is a lower triangular system). Then we solve $\mathbf{L}\mathbf{z} = \mathbf{b}$ (lower triangular system), then $(\mathbf{I} + \mathbf{v}\mathbf{v}^T)\mathbf{w} = \mathbf{z}$ (where we can use a), where we found an expression for $(\mathbf{I}+\mathbf{v}\mathbf{v}^T)^{-1}$), and finally $\mathbf{L}^T\mathbf{x} = \mathbf{w}$ (upper triangular system).

(Continued on page 4.)

Problem 3.

a) Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Explain how we can use the spectral theorem for symmetric matrices to show that

$$\lambda_{min} = \min_{\mathbf{x}\neq 0} R(\mathbf{x}) = \min_{\|\mathbf{x}\|_2=1} R(\mathbf{x}),$$

where λ_{min} is the smallest eigenvalue of $\mathbf{A}$, and $R(\mathbf{x})$ is the Rayleigh quotient given by

$$R(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}.$$

Solution: The spectral theorem says that we can write any real symmetric matrix as $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^T$, where $\mathbf{U}$ is orthogonal and $\mathbf{D}$ is diagonal. We now get that

$$\begin{aligned} R(\mathbf{x}) &= \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \frac{\mathbf{x}^T\mathbf{U}\mathbf{D}\mathbf{U}^T\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \frac{(\mathbf{U}^T\mathbf{x})^T\mathbf{D}(\mathbf{U}^T\mathbf{x})}{\|\mathbf{x}\|^2} \\ &= \frac{(\mathbf{U}^T\mathbf{x})^T\mathbf{D}(\mathbf{U}^T\mathbf{x})}{\|\mathbf{U}^T\mathbf{x}\|^2} = R_D(\mathbf{U}^T\mathbf{x}) \end{aligned}$$

since $\mathbf{U}$ is orthogonal (R_D is the Rayleigh quotient using $\mathbf{D}$ instead of $\mathbf{A}$). We thus have that

$$\min_{\mathbf{x}\neq 0} R(\mathbf{x}) = \min_{\mathbf{x}\neq 0} R_D(\mathbf{U}^T\mathbf{x}) = \min_{\mathbf{x}\neq 0} R_D(\mathbf{x}) = \min_{\mathbf{x}\neq 0} \sum_{i=1}^{n} \lambda_i x_i^2/\|\mathbf{x}\|^2 = \lambda_i,$$

where the minimum is attained for $\mathbf{x} = \mathbf{e}_i$ with $\lambda_i = \lambda_{min}$.

b) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $\|\mathbf{x}\|_2 = 1$ and $\mathbf{y} \neq 0$. Show that

$$R(\mathbf{x} - t\mathbf{y}) = R(\mathbf{x}) - 2t(\mathbf{A}\mathbf{x} - R(\mathbf{x})\mathbf{x})^T\mathbf{y} + \mathcal{O}(t^2).$$

where $t > 0$ is small.

(Hint: Use Taylor's theorem for the function $f(t) = R(\mathbf{x} - t\mathbf{y})$.)

Solution: Using the hint we have that $f(0) = R(\mathbf{x})$. We also have that

$$f(t) = \frac{(\mathbf{x} - t\mathbf{y})^T\mathbf{A}(\mathbf{x} - t\mathbf{y})}{(\mathbf{x} - t\mathbf{y})^T(\mathbf{x} - t\mathbf{y})} = \frac{\mathbf{x}^T\mathbf{A}\mathbf{x} - 2t\mathbf{x}^T\mathbf{A}\mathbf{y} + t^2\mathbf{y}^T\mathbf{A}\mathbf{y}}{\|\mathbf{x}\|^2 - 2t\mathbf{x}^T\mathbf{y} + t^2\|\mathbf{y}\|^2} = \frac{g(t)}{h(t)}.$$

We here have that

$$\begin{aligned} &g(0) = \mathbf{x}^T\mathbf{A}\mathbf{x} \qquad && g'(t) = -2\mathbf{x}^T\mathbf{A}\mathbf{y} + 2t\mathbf{y}^T\mathbf{A}\mathbf{y} \qquad && g'(0) = -2\mathbf{x}^T\mathbf{A}\mathbf{y} \\ &h(0) = \|\mathbf{x}\|^2 = 1 && h'(t) = -2\mathbf{x}^T\mathbf{y} + 2t\|\mathbf{y}\|^2 && h'(0) = -2\mathbf{x}^T\mathbf{y} \end{aligned}$$

We now get that

$$\begin{aligned} f'(0) &= \frac{g'(0)h(0) - g(0)h'(0)}{h(0)^2} = -2\mathbf{x}^T\mathbf{A}\mathbf{y} + 2\mathbf{x}^T\mathbf{y}\mathbf{x}^T\mathbf{A}\mathbf{x} \\ &= -2((\mathbf{A}\mathbf{x})^T\mathbf{y} - R(\mathbf{x})\mathbf{x}^T\mathbf{y}) = -2(\mathbf{A}\mathbf{x} - R(\mathbf{x})\mathbf{x})^T\mathbf{y}. \end{aligned}$$

Clearly the second derivative of f is bounded close to 0, so that $f(t) = f(0) + tf'(0) + \mathcal{O}(t^2)$. Inserting $f(0) = R(\mathbf{x})$ and $f'(0) = -2(\mathbf{A}\mathbf{x} - R(\mathbf{x})\mathbf{x})^T\mathbf{y}$ gives the desired result.

(Continued on page 5.)

c) Based on the characterisation given in 3a) above it is tempting to develop an algorithm for computing λ_{min} by approximating the minimum of $R(\mathbf{x})$ over the unit ball

$$B_1 = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 = 1\}.$$

Assume that $\mathbf{x}^0 \in B_1$ satisfies $\mathbf{A}\mathbf{x}^0 - R(\mathbf{x}^0)\mathbf{x}^0 \neq 0$, i.e. $(R(\mathbf{x}^0), \mathbf{x}^0)$ is not an eigenpair for $\mathbf{A}$. Explain how we can find a vector $\mathbf{x}^1 \in B_1$ such that $R(\mathbf{x}^1) < R(\mathbf{x}^0)$.

Solution: If $\mathbf{A}\mathbf{x}^0 - R(\mathbf{x}^0)\mathbf{x}^0 \neq 0$ we can choose a vector $\mathbf{y}$ so that $(\mathbf{A}\mathbf{x}^0 - R(\mathbf{x}^0)\mathbf{x}^0)^T\mathbf{y} > 0$ ($\mathbf{y}$ can for instance be a vector pointing in the same direction as $\mathbf{A}\mathbf{x}^0 - R(\mathbf{x}^0)\mathbf{x}^0$). But then $-2t(\mathbf{A}\mathbf{x}^0 - R(\mathbf{x}^0)\mathbf{x}^0)^T\mathbf{y} < 0$ (t is assumed to be positive here) and since this term dominates $\mathcal{O}(t^2)$ for small t, we see that $R(\mathbf{x}^0 - t\mathbf{y}) < R(\mathbf{x}^0)$. In other words, we can reduce the Rayleigh quotient by taking a small step from $\mathbf{x}^0$ in the direction of $\mathbf{A}\mathbf{x}^0 - R(\mathbf{x}^0)\mathbf{x}^0$.

Good luck!