Foundations of Computational Math I Exam 1

Take-home Exam

Open Notes, Textbook, Homework Solutions Only

Calculators Allowed

No collaborations with anyone

Due beginning of Class Wednesday, October 26, 2011

Question	Points Possible	Points Awarded
1. Basics	25	
2. Linear operators	25	
3. Floating point	25	
4. Factorization	25	
5. Orthogonal Factorization	25	
Total Points	125	

Name:__________________________________

Alias: __________________________________

to be used when posting anonymous grade list.

Problem 1

(25 points)

Suppose $A \in \mathbb{R}^{m \times n}$ and consider the matrix 2-norm

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$$

1.a

(10 points)

Show that $\|A\|_2 \geq \|A_1\|_2$ where

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},$$

$m = m_1 + m_2$, $A_1 \in \mathbb{R}^{m_1 \times n}$, and $A_2 \in \mathbb{R}^{m_2 \times n}$.

Solution:

Recall that $\forall y \in \mathbb{R}^m$, we have

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \rightarrow \|y\|_2^2 = \|y_1\|_2^2 + \|y_2\|_2^2$$

$$\therefore \forall x \in \mathbb{R}^n \quad \|A_1 x\|_2^2 \leq \|A_1 x\|_2^2 + \|A_2 x\|_2^2$$

It follows that

$$\|A\|_2^2 = \max_{\|x\|_2=1} \|Ax\|_2^2 = \max_{\|x\|_2=1} \{\|A_1 x\|_2^2 + \|A_2 x\|_2^2\} \geq \max_{\|x\|_2=1} \|A_1 x\|_2^2 = \|A_1\|_2^2.$$

We therefore have

$$\|A\|_2^2 \geq \|A_1\|_2^2.$$

You can also prove this by exploiting the fact that there exists $x_1 \in \mathbb{R}^n$ such that

$$\|A_1 x_1\|_2 = \|A_1\|_2$$

Therefore,

$$\|A\|_2^2 \geq \|Ax_1\|_2^2 = \|A_1 x_1\|_2^2 + \|A_1 x_1\|_2^2 \geq \|A_1 x_1\|_2^2 = \|A_1\|_2^2$$

Note this proof does not work if you start with $z \in \mathbb{R}^n$ such that $\|Az\|_2 = \|A\|_2$.

A more elegant proof uses the consistency of the family of matrix 2-norms. Define

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} B = \begin{pmatrix} I_{m_1} & 0 \end{pmatrix}$$

Note that it is easily seen from the definition of the matrix 2-norm and B that

$$\|B\|_2 = 1$$

We therefore have

$$A_1 = BA \rightarrow \|A_1\|_2 = \|BA\|_2 \leq \|B\|_2 \|A\|_2 = \|A\|_2$$

as desired.

1.b

(15 points)

Let $\mathcal{S}_1 \subset \mathbb{R}^n$ and $\mathcal{S}_2 \subset \mathbb{R}^n$ be two subspaces of $\mathbb{R}^n$.

(i) **(5 points)** – Suppose $x_1 \in \mathcal{S}_1$, $x_1 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. $x_2 \in \mathcal{S}_2$, and $x_2 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. Show that x_1 and x_2 are linearly independent. **Solution:**

Proof by contradiction. Assume they are dependent. If they are dependent then we must have $x_1 = \alpha x_2$. $x_1 = \alpha x_2 \rightarrow x_1 \in \mathcal{S}_2$ which is a contradiction. $\therefore x_1,\ x_2$ are linearly independent.

(ii) **(10 points)** – Suppose $x_1 \in \mathcal{S}_1$, $x_1 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. $x_2 \in \mathcal{S}_2$, and $x_2 \notin \mathcal{S}_1 \cap \mathcal{S}_2$. Also, suppose that $x_3 \in \mathcal{S}_1 \cap \mathcal{S}_2$ and $x_3 \neq 0$, i.e., the intersection is not empty. Show that x_1, x_2 and x_3 are linearly independent. (Note the result of the previous part of the problem may be useful.) **Solution:**

Proof by contradiction. Assume they are dependent. Since we know x_1 and x_2 are linearly independent, linear dependence implies $x_3 = \alpha_1 x_1 + \alpha_2 x_2$. Three cases possible:

1. $\alpha_1 \neq 0$ and $\alpha_2 = 0 \rightarrow x_3 \in \mathcal{S}_1\ \ and\ \ x_3 \notin \mathcal{S}_2$. Contradiction.
2. $\alpha_2 \neq 0$ and $\alpha_1 = 0 \rightarrow x_3 \in \mathcal{S}_2\ \ and\ \ x_3 \notin \mathcal{S}_1$. Contradiction.
3. $\alpha_1 \neq 0$ and $\alpha_2 \neq 0 \rightarrow x_2 = \frac{\alpha_1}{\alpha_2} x_1 - \frac{1}{\alpha_2} x_3 \in \mathcal{S}_1$. Contradiction.

$\therefore x_1,\ x_2,\ x_3$ are linearly independent.

Problem 2

(25 points)

2.a

(15 points)

Recall that $\mathcal{P}_n$, the set of polynomials of degree less than or equal to n, and the operation of polynomial addition is equivalent to the vector space $\mathbb{C}^{n+1}$.

i. **(5 points)** Show that the mapping from a polynomial $p(\tau) \in \mathcal{P}_n$ to its derivative with respect to τ, $p'(\tau) \in \mathcal{P}_n$ can be expressed as an $n+1 \times n+1$ matrix applied to a vector v, i.e., $v' = Dv$, where the vector $v \in \mathbb{C}^{n+1}$ represents $p(\tau)$ and the vector $v' \in \mathbb{C}^{n+1}$ represents $p'(\tau)$.

ii. **(5 points)** What is the null space $\mathcal{N}(D)$ and how does it relate to the derivatives of the polynomials?

iii. **(5 points)** Recall that the $n+1-$st derivative of a polynomial of degree less than or equal to n is identically 0. How is this reflected in the algebraic properties of D?

Solution:

The pattern is clear from $\mathcal{P}_3$ and $\mathbb{C}^4$. The polynomial to vector correspondence can be chosen with respect to any polynomial basis. The simplest is the standard monomial basis which yields

$$\begin{array}{rcl} p(\tau) & = & \nu_0 + \nu_1\tau + \nu_2\tau^2 + \nu_3\tau^3 \\ & \Updownarrow & \\ v & = & \begin{pmatrix} \nu_0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \end{array}$$

Differentiation is a simple scale and shift of the coefficients, i.e.,

$$\begin{array}{rcl} v' & = & Dv \\ \begin{pmatrix} \nu_1 \\ 2\nu_2 \\ 3\nu_3 \\ 0 \end{pmatrix} & = & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \nu_0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \end{array}$$

The null space is $\mathcal{N}(D) = span[e_1]$ and this reflects that fact that a constant has a derivative that is identically 0.

Higher order derivatives are simply powers of D and since $D^4 = 0$ we have the desired result that $p''''(\tau) \equiv 0$ for this example. The fact that $D^{n+1} = 0$ for any n can be deduced by from the pattern seen when verifying $D^4 = 0$.

2.b

(10 points)

Consider computing the matrix vector product $y = Tx$, i.e., you are given T and x and you want to compute y. Suppose further that the matrix $T \in \mathbb{R}^{n\times n}$ is tridiagonal with constant values on each diagonal. For example, if $n = 6$ then

$$\begin{pmatrix} \alpha & \beta & 0 & 0 & 0 & 0 \\ \gamma & \alpha & \beta & 0 & 0 & 0 \\ 0 & \gamma & \alpha & \beta & 0 & 0 \\ 0 & 0 & \gamma & \alpha & \beta & 0 \\ 0 & 0 & 0 & \gamma & \alpha & \beta \\ 0 & 0 & 0 & 0 & \gamma & \alpha \end{pmatrix}$$

(i) Write a simple loop-based psuedo-code that computes $y = Tx$ for such a matrix $T \in \mathbb{R}^n$.

(ii) How many operations are required as a function of n?

(iii) How many storage locations are required as a function of n?

Solution:

$$\eta_i \leftarrow \gamma\xi_{i-1} + \alpha\xi_i + \beta\xi_{i+1}, \quad 2 \le i \le n-1$$

$$\eta_1 \leftarrow \alpha\xi_1 + \beta\xi_2$$

$$\eta_n \leftarrow \gamma\xi_{n-1} + \alpha\xi_n$$

The storage required is $2n+3$: n for x, n for y, and 3 for T. The computational complexity is $5(n-2) + 6 = 5n - 4 = 5n + O(1)$.

Problem 3

(25 points)

3.a

(20 points)

Define the function $f(x) = x - 1$ on the domain $x > 1$. Let $x_0 \in \mathbb{R}$, $x_0 > 2$, and $x_1 = x_0(1+\delta)$ where $\delta \in \mathbb{R}$ with $|\delta| < 1$.

(i) **(10 points)** Determine the relative error between $f(x_1)$ and $f(x_0)$, and the relative condition number $\kappa_{rel}(x_0)$.

Solution: We have

$$\frac{|f(x_1) - f(x_0)|}{|f(x_0)|} = \frac{|x_0|}{|x_0 - 1|}|\delta|$$

$$\frac{|f(x_1) - f(x_0)|}{|f(x_0)|} \leq \kappa_{rel}(x_0)|\delta| \rightarrow \kappa_{rel}(x_0) = \frac{|x_0|}{|x_0 - 1|} = \frac{x_0}{x_0 - 1} \quad \text{since } x_0 > 1$$

This is consistent with the Taylor series form. Note for $x_0 > 2$ we should take

$$\kappa_{rel}(x_0) > 1$$

as required.

(ii) **(10 points)** Suppose $|\delta| < 10^{-7}$. Can we expect that the relative error between $f(x_1)$ and $f(x_0)$ is no more than 10^{-4} for the region of values assumed for x_0?

Solution:

We want

$$\frac{|f(x_1) - f(x_0)|}{|f(x_0)|} \leq 10^{-4}$$

and we are given that $|\delta| < 10^{-7}$. Therefore, since

$$\frac{|f(x_1) - f(x_0)|}{|f(x_0)|} \leq \kappa_{rel}|\delta|$$

we must have

$$\kappa_{rel} = \frac{x_0}{x_0 - 1} \leq 10^3$$

We therefore have

$$\frac{x_0}{x_0 - 1} \leq 10^3 \rightarrow x_0 \geq \frac{1000}{999}$$

This is satisfied for all of the region of interest, $x_0 > 2$.

3.b

(5 points)

Suppose x, y and z are floating point numbers in a standard model floating point arithmetic system. Is it true that

$$(x \boxed{op} (y \boxed{op} z)) = ((x \boxed{op} y) \boxed{op} z) \ ?$$

Solution: Floating point arithmetic is not associative. To prove it we need a counterexample. Recall our discussion of cancellation.

$$x = 472635.0000 \;\; y = 27.5013 \;\; z = -472630.0000$$
$$fl(fl(472635.0000 + 27.5013) - 472630) = 33$$
$$fl(27.5013 + fl(472635 - 472630)) = 32.5013$$

Problem 4

(25 points)

4.a

(15 points)

If $A \in \mathbb{R}^{n \times n}$ is a matrix with rank $1 \leq k < n$ then there exists two matrices $X \in \mathbb{R}^{n \times k}$ and $Y \in \mathbb{R}^{n \times k}$ both of which have full column rank k and such that

$$A = XY^T$$

This is called a **full rank factorization** of A.

The reverse is also true, i.e., if there exist two matrices $X \in \mathbb{R}^{n \times k}$ and $Y \in \mathbb{R}^{n \times k}$ both of which have full column rank k such that $A = XY^T$, then the rank of A is k.

(i) **(5 points)** Show that the full rank factorization of A **is not unique**.

(ii) **(5 points)** Find a basis for $\mathcal{R}(A)$, the range of A.

(iii) **(5 points)** Characterize a vector in the null space $\mathcal{N}(A)$.

Solution:

For any nonsingular $M \in \mathbb{R}^{k \times k}$ we have

$$A = XY^T = XMM^{-1}Y^T = (XM)(YM^{-T})^T = \tilde{X}\tilde{Y}^T.$$

To determine the range of A note that $Ac = XY^T c = Xv$. Since Y is rank k so is Y^T and therefore, for any $d \in \mathbb{R}^k$ there is at least one $c \in \mathbb{R}^n$ such that $d = Y^T c$. As a result, we have $\mathcal{R}(A) = \mathcal{R}(X)$.

If $c \in \mathcal{N}(A)$ then $Ac = 0$ and $XY^T c = 0$. Since X is full rank, i.e., linearly independent columns, this implies $Y^T c = 0$. Therefore, $c \perp \mathcal{R}(Y)$ and $\mathcal{R}^{\perp}(Y) = \mathcal{N}(A)$.

4.b

(10 points)

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. Suppose when computing the Cholesky factorization of A using IEEE floating point arithmetic we encounter at some step a computed Schur complement that is identically 0, i.e., every element in the computed Schur complement has the value of 0. What can we conclude about the original matrix A? Justify your answer.

Solution:

The Cholesky factorization in finite precision produces a factorization of a perturbed matrix $A + E$. Since the Schur complement is identically zero after, say k, steps we have

$$A + E = L_k L_k^T$$

where $L_k \in \Re^{n \times k}$ is lower trapezoidal with positive elements on the diagonal. This is a rank factorization of $A + E$ which is therefore a singular matrix with rank $k < n$.

We conclude that the original matrix A is near a singular matrix $A + E$, i.e., it is ill-conditioned.

Problem 5

(25 points)

5.a

(10 points)

Recall, that, given a full column rank matrix $A \in \mathbb{R}^{n\times k}$, we have discussed a reliable algorithm to compute an orthonomal basis of $\mathcal{R}(A)$ by computing the Householder reflectors $H_1^T, \ldots, H_k^T$ that transform A

$$H_k^T \ldots H_1^T A = \begin{pmatrix} R \\ 0 \end{pmatrix}$$

where R is upper triangular and nonsingular, and then evaluating the computing efficiently the first k columns of

$$H = H_1 \ldots H_k$$

to get the k orthonormal columns of Q where $A = QR$.

(i) Suppose you want to compute an orthonormal basis of $\mathcal{R}(A)$ but the H_i were not saved, i.e., we have only R and A. Describe how you would compute Q where $A = QR$. You need not worry about numerical issues.

(ii) How many operations are required as a function of n and k?

Solution:

Given $A = QR$ with A and R known it follows trivially that

$$Q = AR^{-1}$$

This is computed as a set of triangular system solves since

$$A = QR \rightarrow R^T Q^T = A^T$$
$$R^T(Q^T e_i) = A^T e_i$$

So each row of Q is computed by solving a lower triangular linear system with the corresponding row of A as the righthand side vector. Each requires k^2 computations therefore the total complexity is nk^2.

5.b

(15 points)

Suppose you are given the nonsingular tridiagonal matrix $T \in \mathbb{R}^{n\times n}$ For example, if $n = 6$ then

$$\begin{pmatrix} \alpha_1 & \beta_1 & 0 & 0 & 0 & 0 \\ \gamma_2 & \alpha_2 & \beta_2 & 0 & 0 & 0 \\ 0 & \gamma_3 & \alpha_3 & \beta_3 & 0 & 0 \\ 0 & 0 & \gamma_4 & \alpha_4 & \beta_4 & 0 \\ 0 & 0 & 0 & \gamma_5 & \alpha_5 & \beta_5 \\ 0 & 0 & 0 & 0 & \gamma_6 & \alpha_6 \end{pmatrix}$$

(i) Suppose you use Householder reflectors to transform T to upper triangular, i.e.,

$$H_{n-1}^T \dots H_1^T T = R.$$

What is the zero/nonzero structure of R?

(ii) What is the structure of each of the reflectors H_i?

(iii) What is the computational complexity of the factorization, i.e., what is k in $O(n^k)$? (You do not have to determine the constant in the complexity expression.)

Solution:

Consider the form of the matrix

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & 0 & 0 & 0 \\ \gamma_2 & \alpha_2 & \beta_2 & 0 & 0 & 0 \\ 0 & \gamma_3 & \alpha_3 & \beta_3 & 0 & 0 \\ 0 & 0 & \gamma_4 & \alpha_4 & \beta_4 & 0 \\ 0 & 0 & 0 & \gamma_5 & \alpha_5 & \beta_5 \\ 0 & 0 & 0 & 0 & \gamma_6 & \alpha_6 \end{pmatrix}$$

H_1^T eliminates the nonzeros below the diagonal in the first column. There is only γ_2 so this should influence the structure of H_1^T. Applying the definition of the Householder reflector we have

$$v = \begin{pmatrix} \alpha_1 \\ \gamma_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad x = \begin{pmatrix} \alpha_1 \pm \|v\|_2 \\ \gamma_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mu = -\frac{2}{\|x\|_2^2}$$

$$H_1^T = I + \mu x x^T$$

Now consider the zero/nonzero structure of H_1^T where $*$ mark positions where nonzero elements can occur

$$H_1^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} * \\ * \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} * & * & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Since H_1^T is orthogonal the 2×2 block must be orthogonal and have elements that are *sin* and *cos* with the appropriate signs. Note also that when H_1^T is applied to T only the first two rows are modified and we have the pattern

$$H_1^T T = \begin{pmatrix} * & * & * & 0 & 0 & 0 \\ 0 & * & * & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \end{pmatrix}$$

Note the additional nonzero position in the first row.

Similar reasoning yields

$$H_2^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & 0 & 0 & 0 \\ 0 & * & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$H_2^T H_1^T T = \begin{pmatrix} * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \end{pmatrix}$$

We therefore have for $n = 6$

$$H_5^T \cdots H_1^T T = \begin{pmatrix} * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

and

$$H_3^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad H_4^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad H_5^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * & * \end{pmatrix}$$

which demonstrates the patterns in H_i^T and R for any n. That is, H_i^T has a single nontrivial 2×2 orthogonal matrix starting in position (i, i) and R has a main diagonal and two superdiagonals.