University of Colorado Denver
Department of Mathematical and Statistical Sciences
Applied Linear Algebra Ph.D. Preliminary Exam
January 13, 2014

Name: ____

Exam Rules:

- This is a closed book exam. Once the exam begins, you have 4 hours to do your best. Submit as many solutions as you can. All solutions will be graded and your final grade will be based on your six best solutions.
- Each problem is worth 20 points.
- Justify your solutions: **cite theorems that you use**, provide counter-examples for disproof, give explanations, and show calculations for numerical problems.
- If you are asked to prove a theorem, do not merely quote that theorem as your proof; instead, produce an independent proof.
- Begin each solution on a new page and use additional paper, if necessary.
- Write only on one side of paper.
- Write legibly using a dark pencil or pen.
- Ask the proctor if you have any questions.

Good luck!

1.	________	5.	________
2.	________	6.	________
3.	________	7.	________
4.	________	8.	________
		Total	________

DO NOT TURN THE PAGE UNTIL TOLD TO DO SO.

Applied Linear Algebra Preliminary Exam Committee:
Joshua French, Julien Langou (Chair), Anatolii Puhalskii.

1. Let A be a full column rank n–by–k matrix (so $k \leq n$) and b to be a column vector of size n. We want to minimize the squared Euclidean norm $L(x) := ||Ax - b||_2^2$ with respect to x.

   (a) Prove that, if $\text{rank}(A) = k$, then $A^T A$ is invertible.

   (b) Compute the gradient of $L(x)$.

   (c) Directly derive the normal equations by minimizing $L(x)$, and then provide the closed-form expression for x that minimizes $L(x)$.

   (d) We consider a QR factorization of A where Q is n-by-k and R is k-by-k. Show that an equivalent solution for x is $x = R^{-1}Q^T b$.

Solution

(a) Let x such that $A^T Ax = 0$, then $x^T A^T Ax = 0$ so that $\|Ax\|^2 = 0$ so that $Ax = 0$. But, since A is full column rank, $\text{Null}(A) = \{0\}$, so that $Ax = 0 \Rightarrow x = 0$. We proved that $A^T Ax = 0 \Rightarrow x = 0$. Since $A^T A$ is square, this means that $A^T A$ is invertible.

(b) The gradient of $L(x) = (Ax - b)^T (Ax - b) = x^T A^T Ax - 2x^T bA^T b + b^T b$ is $\nabla L(x) = 2A^T Ax - 2A^T b$.

(c) Setting the gradient to zero, we get the normal equations $A^T Ax = A^T b$, by question (a), we know that $A^T A$ is invertible, the unique solution of the normal equations is obtained as $x = (A^T A)^{-1} A^T b$.

(d) The QR factorization of A has the property $A = QR$, with $Q^T Q = I$. (We note that R is upper triangular but this does not matter here.) Starting from the normal equations in (a), we have $R^T Q^T QRx = R^T Q^T b$, which simplifies to $R^T Rx = R^T Q^T b$ since $Q^T Q = I$. We note that, since A has full column rank, this means that R is invertible. (Proof. By contrapositive. Assume R is not invertible, then there exists x nonzero such that $Rx = 0$, so that $QRx = 0$ so that $Ax = 0$ (with x nonzero) so $\dim(\text{Null}(A) > 0$ so $\text{Rank}(A) < k$ so A is not full column rank.) Since R is invertible, (so is R^T,) from $R^T Rx = R^T Q^T b$, we get $x = R^{-1} Q^T b$.

2. Let V be a real vector space.

   (a) Give the definition of a real inner product $\langle \cdot, \cdot \rangle$ over the vector space V. (That is the set of properties from the definition of of a real inner product.)

   We define $\|x\|$ as $\|x\| = \sqrt{\langle x, x \rangle}$.

   (b) From these two definitions, state and prove the Cauchy-Schwarz inequality.

   (c) Now, state and prove the triangular inequality.

   (d) Now, prove that $\|x\|$ is a norm.

   Solution

   (a) A real inner product on V is a funcion from V^2 to $\mathbb{R}$ with the following properties:

      i. for all x in V, $\langle x, x \rangle \geq 0$,

      ii. $\langle x, x \rangle = 0$ if and only if $x = 0$,

      iii. for all x in V, for all y in V, $\langle x, y \rangle = \langle y, x \rangle$,

      iv. for all x in V, for all y in V, for all z in V, $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$,

      v. for all α in $\mathbb{R}$, for all x in V, for all y in V, $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$.

   (b) We note that by property (i) above, for all x in V, $\langle x, x \rangle \geq 0$, and so $\|x\| = \sqrt{\langle x, x \rangle}$ is well defined for x in V.

   The Cauchy-Schwarz inequality states that, for all u and all v, we have

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

   Now we write that

$$\begin{aligned} 0 &\leq \langle \|u\| v - \|v\| u, \|u\| v - \|v\| u \rangle \\ &= \|u\|^2 \langle v, v \rangle - 2\|u\| \|v\| \langle u, v \rangle + \|v\|^2 \langle u, u \rangle \\ &= 2\|u\|^2 \|v\|^2 - 2\|u\| \|v\| \langle u, v \rangle. \end{aligned}$$

   Rearranging yields

$$\begin{aligned} 2\|u\| \|v\| \langle u, v \rangle &\leq 2\|u\|^2 \|v\|^2 \\ \langle u, v \rangle &\leq \|u\| \|v\|. \end{aligned}$$

   We can apply the same reasonning to $-u$ instead of u and we obtain the Cauchy-Schwarz inequality.

   (c) The triangle inequality states that, for all u and all v, we have

$$\|u + v\| \leq \|u\| + \|v\|.$$

Note that

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \text{ by Cauchy-Schwarz inequality} \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

and the inequality follows by taking the square root of both sides.

(d) A norm is a function from V to $\mathbb{R}$ with the following properties:

i. for all x in V, ($\|x\| = 0 \Rightarrow x = 0$),

ii. for all x in V, for all α in $\mathbb{R}$, $\|\alpha x\| = |\alpha|\|x\|$,

iii. for all x in V, for all y in V, $\|x+y\| \leq \|x\| + \|y\|$.

Property (2.d.i) comes from property (2.a.ii). Property (2.d.ii) comes from property (2.a.iii) and property (2.a.v). Property (2.d.iii) is the triangular inequality which we prove in (2.c).

3. Suppose A is a positive definite symmetric real n–by–n matrix and B is a real m–by–n matrix such that BB^T is positive definite. Prove that the matrix $B^T(BA^{-1}B^T)^{-1}B$ is symmetric positive definite.

   Solution

   Since A is positive definite, A^{-1} is positive definite. For $x \in \mathbb{R}^m$, $B^Tx = 0 \in \mathbb{R}^n$ if and only if $x = 0$. (If $B^Tx = 0$ for $x \neq 0$, then $BB^Tx = 0$ which is impossible by BB^T being positive definite.) Hence, $x^TBA^{-1}B^Tx = 0$ if and only if $x = 0$, so $BA^{-1}B^T$ is positive definite. Therefore, $(BA^{-1}B^T)^{-1}$ is positive definite which implies, as before that $B^T(BA^{-1}B^T)^{-1}B$ is positive definite.

4. Suppose A is a positive definite symmetric square real matrix and B is a symmetric square real matrix. Show that there exists a square real matrix C such that $C^T AC$ is the identity matrix and $C^T BC$ is a diagonal matrix.

Solution

Let $C_1 = A^{1/2}$. Then $C_1^{-1} AC_1^{-1}$ is the identity matrix and $C_1^{-1} BC_1^{-1}$ is symmetric. We can write $C_1^{-1} BC_1^{-1} = PDP^T$, where D is diagonal and P is orthogonal. Then $D = (P^T C_1^{-1})B(C_1^{-1} P)$ and $(P^T C_1^{-1})A(C_1^{-1} P) = P^T (C_1^{-1} AC_1^{-1})P$ is the identity matrix. Thus, one can take $C = C_1^{-1} P$.

5. Let $\mathcal{P}_n$ represent the real vector space of polynomials in x of degree less than or equal to n defined on $[0,1]$. Given a real number a, we define $Q_n(a)$ the subset of $\mathcal{P}_n$ of polynomials that have the real number a as a root.

   (a) Let a be a real number. Show that $Q_n(a)$ is a subspace of $\mathcal{P}_n$. Determine the dimension of that subspace and exhibit a basis.

   (b) Let the inner product in $\mathcal{P}_n$ be defined by $\langle p, q\rangle = \int_0^1 p(x)q(x)dx$. Determine the orthogonal complement of the subspace $Q_2(1)$ of $\mathcal{P}_2$.

Solution

(a) Polynomials in $Q_n(a)$ can be written as $p(x) = (x-a)q(x)$ where $q(x)$ is a polynomial of degree less than or equal to $n-1$. The definition of a subspace is verified routinely. Since $Q_n(a)$ is isomorphic with $\mathcal{P}_{n-1}$, its dimension is n, $\{(x-a), (x-a)^2, \ldots, (x-a)^n\}$ is a basis.

(b) We can write a polynomial in $\mathcal{P}_2$ as $a_0 + a_1(x-1) + a_2(x-1)^2$. We need a polynomial orthogonal to $x-1$ and $(x-1)^2$, so

$$\int_0^1 (a_0 + a_1(x-1) + a_2(x-1)^2)(x-1)dx = 0,$$

$$\int_0^1 (a_0 + a_1(x-1) + a_2(x-1)^2)(x-1)^2 dx = 0,$$

which yields

$$-\frac{a_0}{2} + \frac{a_1}{3} - \frac{a_2}{4} = 0,$$
$$\frac{a_0}{3} - \frac{a_1}{4} + \frac{a_2}{5} = 0,$$

so

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = a_2 \begin{pmatrix} 3/10 \\ 6/5 \\ 1 \end{pmatrix}$$

Thus, $Q_2(a)^{\perp} = \{3a_2 + 12a_2(x-1) + 10a_2(x-1)^2, a_2 \in \mathbb{R}\}$.

6. Let $\mathbb{F}$ be a commutative field, let $(V, +, .)$ be a vector space over $\mathbb{F}$, let A and B be two subspaces of V, let A' be a subspace such that $A' \oplus (A \cap B) = A$ and let B' be a subspace such that $B' \oplus (A \cap B) = B$. Show that $A + B = (A \cap B) \oplus A' \oplus B'$.

   Solution

   One can write

$$A + B = (A' + (A \cap B)) + (B' + (A \cap B)) = A' + B' + (A \cap B).$$

   So the real question is not about the sum but about the direct sum of $(A \cap B)$, A', and B'.

   Let $x \in (A \cap B)$, $a' \in A'$, $b' \in B'$ such that

$$x + a' + b' = 0.$$

   Then, on the one hand, $b' \in B'$ but $B' \subset B$, so $b' \in B$, on the other hand, $b' = -x - a'$, but $x \in A$ (since $x \in (A \cap B)$), and $a' \in A$ (since $a' \in A'$ and $A' \subset A$), so $b' \in A$. We see that $b' \in (A \cap B))$. However, we also have that $b' \in B'$. Therefore $b' \in (A \cap B) \cap B'$. But $(A \cap B)$ and B' are in direct sum so $(A \cap B) \cap B' = \{0\}$, so $b' = 0$.

   Now we have

$$x + a' = 0.$$

   $x \in (A \cap B)$, $a' \in A'$, but, since $(A \cap B)$ and A' are in direct sum, $x = 0$ and $a' = 0$.

   We prove that $x = 0$, $a' = 0$, and $b' = 0$. Therefore $(A \cap B)$, A', and B' are in direct sum and

$$A + B = (A \cap B) \oplus A' \oplus B'.$$

7. Let $\mathbb{F}$ be a commutative field, let $(V, +, .)$ be a vector space over $\mathbb{F}$, let n be a natural number, let $(e_1, \ldots, e_n)$ be a linear independent list in V, let $\lambda_1, \ldots, \lambda_n$ be n scalars in $\mathbb{F}$, let $u = \sum_{i=1}^n \lambda_i e_i$, and let, for all $i = 1, \ldots, n$, $v_i = u + e_i$. Show that $(v_1, \ldots v_n)$ is linearly dependent if and only $\sum_{i=1}^n \lambda_i = -1$.

Solution

First, let us that assume $(v_1, \ldots v_n)$ is linearly dependent, then there exists n scalars α_1, ..., α_n, not all zeros such that,

$$\sum_{i=1}^n \alpha_i v_i = 0.$$

Since, for all $i = 1, \ldots, n$, $v_i = u + e_i$, we have

$$\sum_{i=1}^n \alpha_i (u + e_i) = 0.$$

We split the i sum in two sums:

$$(\sum_{i=1}^n \alpha_i u) + (\sum_{i=1}^n \alpha_i e_i) = 0.$$

Now, we use the fact that $u = \sum_{j=1}^n \lambda_j e_j$:

$$(\sum_{i=1}^n \sum_{j=1}^n \alpha_i \lambda_j e_j) + (\sum_{i=1}^n \alpha_i e_i) = 0.$$

Now, we swap the i and the j sum on the left term and change the dummy index i to a j in the right term:

$$(\sum_{j=1}^n \sum_{i=1}^n \alpha_i \lambda_j e_j) + (\sum_{j=1}^n \alpha_j e_j) = 0.$$

We merge the two j sums and factor the e_j term:

$$\sum_{j=1}^n ((\sum_{i=1}^n \alpha_i)\lambda_j + \alpha_j) e_j = 0. \tag{1}$$

The latter expression reads now as a zero linear combination of the e_j. Since the e_j are linear independent, each of the coefficients in the linear combination has to be 0, this writes:

$$(\sum_{i=1}^n \alpha_i)\lambda_j + \alpha_j = 0, \text{ for } j = 1, \ldots, n$$

We can take the sum for $j = 1$ to n of these n expressions and we get:

$$\sum_{j=1}^{n}[(\sum_{i=1}^{n}\alpha_i)\lambda_j + \alpha_j] = 0.$$

We break the sum in two:

$$\sum_{j=1}^{n}[(\sum_{i=1}^{n}\alpha_i)\lambda_j] + \sum_{j=1}^{n}\alpha_j = 0.$$

We factor the $\sum_{i=1}^{n}\alpha_i$ on the left term:

$$(\sum_{i=1}^{n}\alpha_i)(\sum_{j=1}^{n}\lambda_j) + \sum_{j=1}^{n}\alpha_j = 0.$$

We get

$$(\sum_{i=1}^{n}\alpha_i)\left(1 + \sum_{j=1}^{n}\lambda_j\right) = 0. \tag{2}$$

Now we come back to Equation (1), it read

$$\sum_{j=1}^{n}((\sum_{i=1}^{n}\alpha_i)\lambda_j + \alpha_j)e_j = 0.$$

We see that, if $\sum_{i=1}^{n}\alpha_i = 0$, then $\sum_{j=1}^{n}\alpha_j e_j = 0$, which would imply that the e_j are linearly dependent. Therefore, since the e_j are linearly independent, we have that $\sum_{i=1}^{n}\alpha_i \neq 0$. Now we see that $\sum_{i=1}^{n}\alpha_i \neq 0$ and Equation (2) implies

$$\sum_{j=1}^{n}\lambda_j = -1.$$

This proves that, if $(v_1, \ldots v_n)$ is linearly dependent, then $\sum_{j=1}^{n}\lambda_j = -1$.

Now, let us assume that $\sum_{j=1}^{n}\lambda_j = -1$. We want to prove that $(v_1, \ldots v_n)$ is linearly dependent. That is, we want to find α_i, $i = 1, \ldots, n$, not all zeros, such that

$$\sum_{i=1}^{n}\alpha_i v_i = 0.$$

We will prove that a correct choice for the α_i is $\alpha_i = \lambda_i$. First note that the λ_i are not all zeros since $\sum_{i=1}^{n} \lambda_i = -1$. Second:

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i v_i &= \sum_{i=1}^{n} \lambda_i (u + e_i), \\
&= \sum_{i=1}^{n} (\lambda_i u) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= \sum_{i=1}^{n} (\lambda_i (\sum_{j=1}^{n} \lambda_j e_j)) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} (\lambda_i \lambda_j e_j) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} (\lambda_i \lambda_j e_j) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= \sum_{j=1}^{n} \left((\sum_{i=1}^{n} \lambda_i) \lambda_j e_j \right) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= (\sum_{i=1}^{n} \lambda_i) \sum_{j=1}^{n} (\lambda_j e_j) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= (-1) \sum_{j=1}^{n} (\lambda_j e_j) + \sum_{i=1}^{n} (\lambda_i e_i), \\
&= 0.
\end{aligned}$$

This proves that $(v_1, \dots v_n)$ is linearly dependent.

8. What is the rank of

$$\begin{pmatrix} 1 & a & 1 & b \\ a & 1 & b & 1 \\ 1 & b & 1 & a \\ b & 1 & a & 1 \end{pmatrix}?$$

The rank is a function of a and b. You need to give the values of the rank for all values of $(a, b) \in \mathbb{R}^2$.

Solution

We perform some Gaussian elimination steps.

First, $L_2 \leftarrow L_2 - aL_1$, $L_3 \leftarrow L_3 - L_1$, $L_4 \leftarrow L_4 - bL_1$ gives

$$\begin{pmatrix} 1 & a & 1 & b \\ 0 & 1-a^2 & b-a & 1-ab \\ 0 & b-a & 0 & a-b \\ 0 & 1-ab & a-b & 1-b^2 \end{pmatrix}$$

We assume $a \neq b$ so that we can simplify the third row with $L_3 \leftarrow L_3/(b-a)$, after this we swap second and third row $L_2 \leftrightarrow L_3$. This gives:

$$\begin{pmatrix} 1 & a & 1 & b \\ 0 & 1 & 0 & -1 \\ 0 & 1-a^2 & b-a & 1-ab \\ 0 & 1-ab & a-b & 1-b^2 \end{pmatrix}$$

Now, $L_3 \leftarrow L_3 - (1-a^2)L_1$, $L_4 \leftarrow L_4 - (1-ab)L_1$, gives

$$\begin{pmatrix} 1 & a & 1 & b \\ 0 & 1 & 0 & -1 \\ 0 & 0 & b-a & 2-a^2-ab \\ 0 & 0 & a-b & 2-b^2-ab \end{pmatrix}$$

Finally $L_4 \leftarrow L_4 + L_3$, gives

$$\begin{pmatrix} 1 & a & 1 & b \\ 0 & 1 & 0 & -1 \\ 0 & 0 & b-a & 2-a^2-ab \\ 0 & 0 & 0 & 4-(a+b)^2 \end{pmatrix}$$

So we see that (1) if $a \neq b$ and $a + b \neq \pm 2$, then the rank is 4. (2) if $a \neq b$, and $a + b = \pm 2$, then the rank is 3.

Now let us see to the case when $a = b$. In this case, the matrix is:

$$\begin{pmatrix} 1 & a & 1 & a \\ a & 1 & a & 1 \\ 1 & a & 1 & a \\ a & 1 & a & 1 \end{pmatrix}.$$

It is clear that if $a = 1$ then the rank is 1, if $a \neq 1$, the rank is 2.

Let us repeat:

(a) If $a = b = 1$, then the rank is 1,

(b) If $a = b$ and $a \neq 1$, then the rank is 2,

(c) If $a \neq b$ and $a + b = \pm 2$, then the rank is 3,

(d) If $a \neq b$ and $a + b \neq \pm 2$, then the rank is 4.