HOMEWORK 2, STAT 961: Lin. Alg. 2

Due Fri, 2019/10/11, 5:00pm

Your Name: ... (replace this)

October 3, 2019

Instructions: Edit this LaTex file by inserting your solutions after each problem statement. Generate a PDF file from it and e-mail the PDF in an attachment with filename

hw02-Yourlastname-Yourfirstname.pdf

to the following class Gmail address:

stat961.at.wharton@gmail.com

with **subject line exactly** as follows for easy gmail search:

Homework 2, 2019

Rules to be strictly followed under honor code:

1. You must write your own solutions and not copy from anyone. Verbatim copying from others or unlisted sources will result in zero points for the whole homework.
2. Subject to the previous item, you may explain the problems, but not the solutions, to each other in general terms.
3. Do not discuss the homework with previous years' students of Stat 961/541, and do not consult solutions of similar homeworks of previous years.
4. Report here (1) from whom you got help, (2) who you helped, and (3) other sources, such as online, papers, books, ... You do not need to report help with LaTex and English language.
   - **Who helped me:** ... (replace this)

- **Who I helped:** ... (replace this)
- **Complete list of my other sources:** ... (replace this)

Instructions for presentation and typesetting:

1. Give derivations where appropriate, but don't when instructed to give the answer without derivation.
2. Notations are different here from HW1. The reason is that we need to distinguish between abstract objects and their coordinates. We use bold for abstract objects and italic for coordinate objects.
3. Matrix transposes and inverses should be written as $M^\top$ and M^{-1}; simultaneous transposing and inverting can be written $M^{-\top}$.
4. To mimic obvious R functions in LaTex math mode, use cbind(...), rbind(...).

Problems for Review of Abstract Linear Algebra

Homework 2, Stat 961, 2019C

1. What follows is remedial material about vector spaces, bases, coordinates, linear forms, linear maps, bilinear forms, basis changes and associated coordinate transformations.

   - Let $\mathbb{V}$ and $\mathbb{W}$ be linear spaces (synonym: vector spaces) over the real numbers, that is, addition of elements and multiplication of elements with real numbers are defined and follow the usual axioms.
   - The linear spaces $\mathbb{V}$ and $\mathbb{W}$ are "abstract," meaning that you should not think of them as spaces of coordinate tuples. Anything we do should apply to all kinds of concrete linear spaces, such as spaces of polynomials, splines, signed measures (of measure theory), tensors, tangent vectors at a point on a curved manifold, etc.
   - However, one of the purposes here is to introduce the notion of a basis and coordinates relative to a basis. Subsequently we do the same for linear forms, $\mathbb{V} \to \mathbb{R}$, that is, linear functions on $\mathbb{V}$ (also called "dual vectors"), then for linear maps $\mathbb{V} \to \mathbb{W}$, and finally for bilinear forms $\mathbb{V} \times \mathbb{W} \to \mathbb{R}$, which are linear in each of two vector arguments separately.
   - In Problem 2 we consider the effects of a change of basis on the coordinates of all four objects: vectors, linear forms, linear maps, and bilinear forms.
   - In Problem 3 we apply the concepts to concrete spaces of polynomials. You will need to apply the concepts of Problems 1 and 2 rigorously to get it right.
   - Important: There is no inner product yet, hence there are no notions of length, angle, orthogonality, symmetry of linear maps, and identification of primal and dual vectors. This will require careful thinking because we are used to using these concepts, often without realizing it.
   - **Definitions**:
     - A finite set of vectors $\{\mathbf{b}_1, ..., \mathbf{b}_p\} \subset \mathbb{V}$ is said to be **linearly independent** if $\sum_{j=1...p} x_j \mathbf{b}_j = \mathbf{0}$ entails $x_j = 0$ for all $j = 1...p$. These vectors are necessarily non-zero ($\neq \mathbf{0}$).

- If $\{\mathbf{b}_1, ..., \mathbf{b}_p\}$ is a maximal set of linearly independent vectors, that is, if increasing the set by any vector $\mathbf{b}_{p+1}$ makes the $p+1$ vectors linearly dependent, then this set is called **a basis of** $\mathbb{V}$.
- Fact: Any other basis must also be of size p. This unique number p is called **the dimension of** $\mathbb{V}$: $p = \dim(\mathbb{V})$. We assume $p > 0$ to avoid dealing with $\mathbb{V} = \{\mathbf{0}\}$.
- If we need another linear space besides $\mathbb{V}$, we denote it by $\mathbb{W}$ and its dimension by $n = \dim(\mathbb{W})$. A basis of $\mathbb{W}$ will be written $\{\mathbf{c}_1, ..., \mathbf{c}_n\} \subset \mathbb{W}$.

[Remarks: (1) We purposely wrote "a basis" and not "the basis" because there always exists an infinity of different bases. (2) There exist infinite-dimensional linear spaces, and you study them in a course on "real analysis" or "functional analysis."]

(a) A vector $\mathbf{x} \in \mathbb{V}$ has a unique representation $\mathbf{x} = \sum_{j=1,...,p} x_j \mathbf{b}_j$. We call the numbers $x_1, ..., x_p$ the *coordinates of* $\mathbf{x}$ *in the basis* $\{\mathbf{b}_1, ..., \mathbf{b}_p\}$. We collect these numbers in a $p \times 1$ matrix or column vector $x = (x_1, ..., x_p)^\top \in \mathbb{R}^{p \times 1}$.

Similarly for $\mathbf{y} \in \mathbb{W}$ we will have $\mathbf{y} = \sum_{i=1,...,n} y_i \, \mathbf{c}_i$ with coordinates collected in the $n \times 1$ matrix or column vector $y = (y_1, ..., y_n)^\top \in \mathbb{R}^{n \times 1}$.

Important: $\mathbf{x}$ and x are *not* the same, $\mathbf{y}$ and y are *not* the same, $\mathbb{V}$ is *not* $\mathbb{R}^{p \times 1}$, and $\mathbb{W}$ is *not* $\mathbb{R}^{n \times 1}$.

Comprehension question: For fixed j_0, what is the coordinate vector of $\mathbf{x} = \mathbf{b}_{j_0}$? Reason?

Answer:

(b) The set of linear forms $\mathbf{x} \mapsto \boldsymbol{l}(\mathbf{x})$, $\mathbb{V} \to \mathbb{R}$, is called the "dual space" $\mathbb{V}'$ of $\mathbb{V}$. It is also a linear space of dimension p. Linear forms are sometimes called "dual vectors," so vectors $\mathbf{x} \in \mathbb{V}$ may be called "primal vectors."

Coefficients for a linear form $\boldsymbol{l}$ relative to the primal basis of $\mathbf{b}_j$ can be defined by $l_j = \boldsymbol{l}(\mathbf{b}_j)$. We collect them in a $1 \times p$ matrix or row vector $l = (l_1, ..., l_p) \in \mathbb{R}^{1 \times p}$. We will see below that these coefficients can be interpreted as coordinates in a suitable basis of $\mathbb{V}'$.

Important: $\mathbb{V}'$ is *not* the same as $\mathbb{V}$, and neither is $\mathbb{V}'$ the same as $\mathbb{R}^{1 \times p}$.

Task: Express $\boldsymbol{l}(\mathbf{x})$ in terms of the associated coefficient vector l, the coordinate vector x, and matrix multiplication. Be careful regarding transposing or not. Give a derivation. (The result will justify calling l_j the "coefficients" of $\boldsymbol{l}$.)

Answer:

(c) Why is the formula in the previous question *not* an inner product?

Answer:

(d) For fixed $j_0 \in \{1, ..., p\}$, consider the "coordinate picker" linear form $\boldsymbol{l}(\mathbf{x}) = x_{j_0}$ that picks the j_0'th coordinate of $\mathbf{x}$. What is the coefficient vector l of this linear form? Reason?

Answer:

(e) Denote the "coordinate picker" linear forms by $\boldsymbol{l}_1, ..., \boldsymbol{l}_p : \boldsymbol{l}_j(\mathbf{x}) = x_j$. Caution: Do *not* confuse $\boldsymbol{l}_j$ and l_j!

Task: Write an *arbitrary* linear form $\boldsymbol{l}$ with coefficient vector $l = (l_1, ..., l_p)$ as a linear combination of these coordinate pickers.

Answer:

Implication: The linear forms $\{\boldsymbol{l}_1, ..., \boldsymbol{l}_p\}$ constitute a basis of $\mathbb{V}'$, also called the "dual basis." The coefficients l_j can now be interpreted as coordinates of the linear form $\boldsymbol{l}$ in this basis. The "coordinate picker" basis of $\mathbb{V}'$ depends on the coordinates of $\mathbf{x}$ in $\mathbb{V}$, which in turn depends on the primal basis $\{\mathbf{b}_1, ..., \mathbf{b}_p\}$, hence the dual basis of $\mathbb{V}'$ fully depends on the primal basis of $\mathbb{V}$. Recall that l_j was defined as $l_j = \boldsymbol{l}(\mathbf{b}_j)$.

(f) Someone forms $x + l^\top$. What is its meaning, if any?

Answer:

(g) A linear map $\mathbf{x} \mapsto \mathbf{y} = \mathbf{A}(\mathbf{x})$, $\mathbb{V} \to \mathbb{W}$, has an associated $n \times p$ matrix $A = (A_{ij})$ relative to the bases $\{\mathbf{b}_1, ..., \mathbf{b}_p\} \subset \mathbb{V}$ and $\{\mathbf{c}_1, ..., \mathbf{c}_n\} \subset \mathbb{W}$. The elements A_{ij} of the matrix are defined by $\mathbf{A}(\mathbf{b}_j) = \sum_{i=1,...,n} A_{ij}\mathbf{c}_i$. Be careful about the order of the indices!

Comprehension question: What is the meaning of $(A_{1j}, ..., A_{nj})^\top$? Why is this even a definition? We didn't write "$A_{ij} = ...$," right?

Answer:

(h) Express a linear map $\mathbf{y} = \mathbf{A}(\mathbf{x})$ in terms of the associated matrix A and column vectors $y = (y_1, ..., y_n)^\top$ and $x = (x_1, ..., x_p)^\top$. Give a pedantic derivation.

Answer:

(i) The linear map $\mathbf{A} : \mathbb{V} \to \mathbb{W}$ has an associated dual map $\mathbf{A}' : \boldsymbol{k} \mapsto \boldsymbol{l} = \mathbf{A}'(\boldsymbol{k})$, in the opposite direction, $\mathbb{W}' \to \mathbb{V}'$, defined by composition of $\boldsymbol{k}$ and $\mathbf{A}$: $\mathbf{A}'(\boldsymbol{k}) = \boldsymbol{k} \circ \mathbf{A}$. That is, $(\mathbf{A}'(\boldsymbol{k}))(\mathbf{x}) = \boldsymbol{k}(\mathbf{A}(\mathbf{x}))$ for all $\mathbf{x} \in \mathbb{V}$ and all $\boldsymbol{k} \in \mathbb{W}'$.

Express the dual map $\mathbf{A}'$ in terms of the matrix A, the row vector $k = (k_1, ..., k_n)$ with coefficients for $\boldsymbol{k}$, and $l = (l_1, ..., l_p)$ for $\boldsymbol{l}$. Give a pedantic derivation for l_j.

Answer:

(j) It might bother you that the dual linear map $\mathbf{A}'$ is in the opposite direction from $\mathbf{A}$. Based on the matrix expression you just derived, you might be tempted to invert the direction and express k as a function of l, thereby making $\mathbf{A}$ and $\mathbf{A}'$ more parallel. Why is this not a good idea? What additional assumption would you have to make?

Answer:

(k) In some texts you might read that $A^\top$ is the natural matrix for $\mathbf{A}'$. What is their convention for the dual coordinate vectors?

Answer:

(l) A bilinear form $(\mathbf{y}, \mathbf{x}) \mapsto \mathbf{B}(\mathbf{y}, \mathbf{x})$, $\mathbb{W} \times \mathbb{V} \to \mathbb{R}$, has an associated matrix $B = (B_{ij})$ in the bases $\{\mathbf{c}_1, ..., \mathbf{c}_n\} \subset \mathbb{W}$ and $\{\mathbf{b}_1, ..., \mathbf{b}_p\} \subset \mathbb{V}$, with B_{ij} defined by $B_{ij} = \mathbf{B}(\mathbf{c}_i, \mathbf{b}_j)$.

Comprehension question 1: For fixed $\mathbf{y} \in \mathbb{W}$, what kind of object is the function $\mathbf{x} \mapsto \mathbf{B}(\mathbf{y}, \mathbf{x})$? Similarly, for fixed $\mathbf{x} \in \mathbb{V}$, what is $\mathbf{y} \mapsto \mathbf{B}(\mathbf{y}, \mathbf{x})$?

Answer:

Comprehension question 2: Does the matrix B stand for a linear map $\mathbb{V} \to \mathbb{W}$?

Answer:

(m) Express a bilinear form $\mathbf{B}(\mathbf{y}, \mathbf{x})$ in terms of the associated matrix B and coordinate vectors y and x. Give a pedantic derivation.

Answer:

(n) Weird: Consider the special case $\mathbb{W} = \mathbb{V}'$ and $\mathbf{B}(\boldsymbol{l}, \mathbf{x}) = \boldsymbol{l}(\mathbf{x})$. Does this qualify as a bilinear form? What would be its matrix in terms of primal and dual bases?

Answer:

(o) For $n = \dim(\mathbb{W})$, let $\boldsymbol{l}^{(1)}, \boldsymbol{l}^{(2)}, ..., \boldsymbol{l}^{(n)}$ be a set of n arbitrary linear forms on $\mathbb{V}$ with associated row vectors $l^{(1)}, l^{(2)}, ..., l^{(n)}$. Construct the following map: $\mathbf{C}(\mathbf{x}) = \sum_{i=1...n} \boldsymbol{l}^{(i)}(\mathbf{x})\mathbf{c}_i$. What kind of map is $\mathbf{C}(\cdot)$? Can you give a matrix description?

Answer:

(p) For m unrelated to n or p, let $\boldsymbol{l}^{(1)}, \boldsymbol{l}^{(2)}, ..., \boldsymbol{l}^{(m)}$ be m arbitrary linear forms on $\mathbb{V}$ with associated row vectors $l^{(1)}, l^{(2)}, ..., l^{(m)}$. Similarly, let $\boldsymbol{k}^{(1)}, \boldsymbol{k}^{(2)}, ..., \boldsymbol{k}^{(m)}$ be m linear forms on $\mathbb{W}$ with row vectors $k^{(1)}, k^{(2)}, ..., k^{(m)}$. Construct the following function: $\mathbf{D}(\mathbf{y}, \mathbf{x}) = \sum_{i=1...m} \boldsymbol{k}^{(i)}(\mathbf{y})\boldsymbol{l}^{(i)}(\mathbf{x})$. What kind of function is $\mathbf{D}(\cdot, \cdot)$? Can you give a matrix description?

Answer:

(q) Consider linear regression as in class: We write the rows of the regressor matrix as column vectors $\boldsymbol{x}$ and the coefficient vector as a column vector $\hat{\boldsymbol{\beta}}$. Do they belong to the same vector space? Give an interpretation in light of what we have learned so far. To reason about these vectors, it may help to think like a physicist again. The two types of vectors are, however, somehow connected. How?

Answer:

2. **Basis changes and associated coordinate transformations:** Consider two bases of the same space $\mathbb{V}$: $\mathbf{b}_1^{old}, ..., \mathbf{b}_p^{old}$ and $\mathbf{b}_1^{new}, ..., \mathbf{b}_p^{new}$. For any vector $\mathbf{x}$ we have coordinates in both bases:

$$\mathbf{x} = \sum_j x_j^{new} \mathbf{b}_j^{new} = \sum_{j'} x_{j'}^{old} \mathbf{b}_{j'}^{old}$$

Denote the respective **coordinate** vectors by $x^{new} = (x_1^{new}, ...x_p^{new})^\top$ and $x^{old} = (x_1^{old}, ...x_p^{old})^\top$. To link the two types of coordinates to each other, we express **the old basis in terms of the new basis**, in this order:

$$\mathbf{b}_{j'}^{old} = \sum_j T_{jj'} \mathbf{b}_j^{new} \tag{1}$$

Collect the coefficients in a $p \times p$ matrix $T = (T_{jj'})$. It is important to keep awareness of the convention regarding the order of the subscripts of T.

(a) What does T contain in column j' ?

Answer:

(b) How are the **coordinate vectors** x^{old} and x^{new} for the abstract vector $\mathbf{x}$ related to each other in terms of the matrix T? Give a pedantic derivation.

Answer:

(c) Argue that T should not be viewed as the matrix of a linear map.

Answer:

(d) Can T be singular (rank-deficient)? Yes or no? No derivations but one English sentence for explanation.

Answer:

(e) What is most likely wrong with the operation $x^{old} + x^{new}$?

Answer:

(f) For a linear form $\boldsymbol{l}$ represented by coordinate vectors l^{old} and l^{new} in the respective dual bases, how are l^{old} and l^{new} related to each other? Give a pedantic derivation. Express l^{new} in terms of l^{old}, in this order, and explain why this is natural, unlike the case of dual linear maps.

Answer:

(g) In some texts, the matrix for coordinate transformation of dual vectors is shown as $T^{-\top}$. What is their convention?

Answer:

Preparation for linear maps and bilinear forms: We now need a second space, $\mathbb{W}$, and two bases in it as well. Denote them $\{\mathbf{c}_1^{old}, ..., \mathbf{c}_n^{old}\}$ and $\{\mathbf{c}_1^{new}, ..., \mathbf{c}_n^{new}\}$. Instead of introducing a new symbol for the associated transformation matrix, we subscript both matrices by their spaces: $T_{\mathbb{V}}$ and $T_{\mathbb{W}}$.

(h) For a linear map, $\mathbb{V} \to \mathbb{W}$, $\mathbf{x} \mapsto A(\mathbf{x}) = \mathbf{y}$, represented by the matrix A^{old} in the old bases $\{\mathbf{b}_1^{old}, ..., \mathbf{b}_p^{old}\} \subset \mathbb{V}$ and $\{\mathbf{c}_1^{old}, ..., \mathbf{c}_n^{old}\} \subset \mathbb{W}$, and the matrix A^{new} in the new bases $\{\mathbf{b}_1^{new}, ..., \mathbf{b}_p^{new}\} \subset \mathbb{V}$ and $\{\mathbf{c}_1^{new}, ..., \mathbf{c}_n^{new}\} \subset \mathbb{W}$, how are A^{old} and A^{new} related to each other? Express A^{new} in terms of A^{old}.

Answer:

(i) For a bilinear form $\mathbb{W} \times \mathbb{V} \to \mathbb{R}$, $(\mathbf{y}, \mathbf{x}) \mapsto \mathbf{B}(\mathbf{y}, \mathbf{x})$, represented by matrices B^{old} and B^{new} in the respective bases, how are B^{old} and B^{new} related to each other? Express B^{new} in terms of B^{old}.

Answer:

3. **Linear Algebra in Action: Spaces of Polynomials**

Let the space $\mathcal{P}$ be the polynomials up to degree p:

$$\mathcal{P} = \left\{ f_a(z) = \sum_{j=0,...,p} a_j z^j \;:\; a = (a_0, ..., a_p)^\top \in \mathbb{R}^{(p+1)\times 1} \right\}$$

This is a linear space of dimension $p+1$. The coefficients a_j ($j = 0, 1, ..., p$) form natural coordinates, collected in a column vector $a \in \mathbb{R}^{(p+1)\times 1}$, in the basis of monomials $1, z, z^2, \;...,\; z^p$.

Convention: In this exercise it is convenient to let all indices start at 0 and end at p, hence $j,\, j' = 0, 1, 2, ..., p$, and **not** $1, 2, 3, ..., p+1$.

(a) Consider the linear form on $\mathbb{V} = \mathcal{P}$ generated by expectation, $f_a(z) \mapsto \mathbf{E}[f_a(Z)]$, where $Z \sim N(0, 1)$. What are the coordinates/coefficients of this linear form? You may consult Wikipedia or another online source and use their notations in your answer. Write the row vector l of coordinates of the linear form explicitly for polynomials of degree 8 ($p = 8$) using actual numbers.

Answer:

(b) Consider the linear map generated by differentiation, $\mathcal{P} \to \mathcal{P}$, $f_a(z) \mapsto \frac{d}{dz} f_a(z)$ (hence $\mathbb{V} = \mathbb{W} = \mathcal{P}$). What is its matrix A? Describe its elements $A_{jj'}$ $(j, j' = 0, 1, ..., p)$

and show the matrix with explicit numbers for $p = 4$. Is this linear map non-singular/invertible? What is its rank?

Answer:

(c) Consider the following bilinear form $(f(z), g(z)) \mapsto \mathbf{E}[f(Z)g(Z)]$, $\mathcal{P} \times \mathcal{P} \to \mathbb{R}$, where again $Z \sim N(0,1)$ and $\mathbb{W} = \mathbb{V} = \mathcal{P}$. Show the matrix for $p = 4$. You may re-use what you learned earlier about the linear form $f(z) \mapsto \mathbf{E}[f(Z)]$.

Answer:

(d) Consider a shift of the z-axis, $\tilde{z} = z - t$, and accordingly a shift of the polynomials:

$$f_a(z) = f_a(\tilde{z} + t) = \tilde{f}_{\tilde{a}}(\tilde{z}).$$

The substitution $z = \tilde{z} + t$ produces another polynomial of the same degree but with argument $\tilde{z}$ and new coefficient vector $\tilde{a}$. The substitution amounts to a change of basis with a coordinate transformation described by a matrix T. Find the elements $T_{jj'}$ of this matrix by expressing the old basis in terms of the new basis.
Hint: You will need binomial coefficients $\binom{\cdots}{\cdots}$.
Show the matrix for $p = 4$ and $t = 1$ with actual numbers.

Answer: