

## Homework 2, Stat 541: Linear Algebra, Due Fri, Sept 25, 2009, 12 Noon

Your Name: (replace this with your name)

October 5, 2009

Instructions: Edit this LaTeX file by inserting your solutions after each problem statement. Generate a PDF file from it and e-mail the PDF to the usual class gmail address. You should not just answer the questions but also give evidence or even proof. You can discuss the homework with each other in general terms, but not with previous years' students of Stat 541. Also, you must write your own solutions and not copy from anyone.

1. Interpretations of matrix multiplication: Assume  $\mathbf{X}$  is of size  $n \times p$  and  $\mathbf{B}$  of size  $p \times q$ .

(a) If  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p)$ ,  $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p)^T$ , express

$$\mathbf{XB} = \sum_{j=1, \dots, p} \mathbf{x}_j \mathbf{b}_j^T$$

Interpret the summands: Size? Rank? How would you compute one in R? **A:** The term  $\mathbf{x}_j \mathbf{b}_j^T$  is called a rank-one matrix. It is what the name says: a matrix of rank one. In fact, every rank-one matrix can be written this way. It is the “outer product” of  $\mathbf{x}_j$  and  $\mathbf{b}_j$ ; in R it would be computed by `outer(X[,j], B[j,])`. The above formula says that every matrix can be written as a sum of rank-one matrices in an obvious way: the rank-one summands are the “outer products” of the columns of the first factor and the rows of the second factor.

(b) If  $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$ ,  $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_q)$ , express

$$(\mathbf{XB})_{i,k} = \mathbf{x}_i^T \mathbf{b}_k$$

Interpret these terms. What would it mean if all of them were zero? How is a term computed in R? **A:** The term  $\mathbf{x}_i^T \mathbf{b}_k$  is the inner product of the  $i$ 'th row of  $\mathbf{X}$  and the  $k$ 'th column of  $\mathbf{B}$ . In R it would be computed by `sum(X[i,]*B[,k])`. The above formula says that a matrix product is the collection of inner products of all rows of  $\mathbf{X}$  and all columns of  $\mathbf{B}$ . If they all were zero, it would mean the rows of  $\mathbf{X}$  and the columns of  $\mathbf{B}$  are orthogonal.

2. Write the Euclidean inner product as  $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$  and the Euclidean squared norm as  $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$ . What is the interpretation of

$$\mathbf{y} \mapsto \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \mathbf{x} ?$$

**A:** This is the orthogonal projection of  $\mathbf{y}$  on the direction given by  $\mathbf{x}$ . The scalar factor written as  $\langle \mathbf{y}, \mathbf{x} \rangle / \|\mathbf{x}\|^2 = \sum y_i x_i / \sum x_i^2$  should remind you of the formula for the slope of the simple linear regression coefficient when regressing the response  $\mathbf{y}$  on the predictor  $\mathbf{x}$ ; in fact, this is the slope formula when there is no intercept term. On the side, note that replacing  $\mathbf{x}$  with any non-zero multiple  $c\mathbf{x}$  produces the same mapping ( $c$  cancels out), hence the mapping really depends only on the “direction” of  $\mathbf{x}$ , not its length or sign. — How do you see that this is the orthogonal projection? You can prove that the “residual vector”

$$\mathbf{y} - \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \mathbf{x}$$

is orthogonal to  $\mathbf{x}$  (it's not difficult).

3. If  $\mathbf{A}$  is a  $n \times n$  matrix, when does it satisfy  $\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle$  for all  $\mathbf{x}$  and all  $\mathbf{y}$ ?

**A:**  $\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = (\mathbf{A}\mathbf{x})^T \mathbf{y} = \mathbf{x}^T \mathbf{A}^T \mathbf{y}$  and  $\langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \mathbf{x}^T \mathbf{A}\mathbf{y}$ , hence we have equality iff (“if and only if”)  $\mathbf{A}^T = \mathbf{A}$ , that is, if  $\mathbf{A}$  is symmetric. — Alternatively, you can plug in all basis vectors  $\mathbf{x} = \mathbf{e}_i$  and  $\mathbf{y} = \mathbf{e}_j$  to see that  $A_{ji} = A_{ij}$ .

4. If  $\|\mathbf{x}\|^2 = 1$ , what is the interpretation of  $\mathbf{P} = \mathbf{x}\mathbf{x}^T$ ? In other words, what does  $\mathbf{y} \mapsto \mathbf{P}\mathbf{y}$  do? (Assume the vectors are compatible:  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ .)

**A:** Apply  $\mathbf{P}$  to a vector  $\mathbf{y}$  to see what the linear mapping with  $\mathbf{P}$  does:  $\mathbf{P}\mathbf{y} = (\mathbf{x}\mathbf{x}^T)\mathbf{y} = \mathbf{x}(\mathbf{x}^T \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle \mathbf{x}$ , which is the same mapping as in Problem 2 when  $\mathbf{x}$  is a unit vector.

5. Same question if  $\|\mathbf{x}\|^2 \neq 1$ ?

**A:** If we split the mapping into two parts:

$$\mathbf{P} = \mathbf{xx}^T = \|\mathbf{x}\|^2 \left( \frac{\mathbf{x}}{\|\mathbf{x}\|} \right) \left( \frac{\mathbf{x}}{\|\mathbf{x}\|} \right)^T,$$

we see that  $\mathbf{P}$  consists of an orthogonal projection onto the direction of  $\mathbf{x}$  followed by a “stretch” by a factor  $\|\mathbf{x}\|^2$ . (Note that the vectors inside the big parens are unit vectors.)

6. Assuming  $\|\mathbf{x}\|^2 = 1$ , verify that  $\mathbf{P} = \mathbf{xx}^T$  satisfies (1)  $\mathbf{PP} = \mathbf{P}$  and (2)  $\mathbf{P}^T = \mathbf{P}$ . Property (1) is called idempotence, (2) is called symmetry.

**A:** (1)  $\mathbf{PP} = (\mathbf{xx}^T)(\mathbf{xx}^T) = \mathbf{x}(\mathbf{x}^T\mathbf{x})\mathbf{x}^T = \mathbf{xx}^T = \mathbf{P}$  because  $\mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2 = 1$ . The only “difficulty” here is to “trust” that matrix multiplication is associative, and to see that in the reshuffling of parens a funny thing happens: the inner product collapses to a number, 1, and hence disappears.

$$(2) \mathbf{P}^T = (\mathbf{xx}^T)^T = (\mathbf{x}^T)^T \mathbf{x}^T = \mathbf{xx}^T = \mathbf{P}.$$

7. Assuming  $\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 = 1$ , under what condition is  $\mathbf{P} = \mathbf{xx}^T + \mathbf{yy}^T$  also idempotent? Is it always symmetric?

**A:** By now we should realize that we are adding up two orthogonal projections to form  $\mathbf{P}$ . The trivial example  $\mathbf{x} = \mathbf{y}$  shows right away that idempotence cannot be generally true because in this case  $\mathbf{P}$  amounts to a projection onto  $\mathbf{x}(= \mathbf{y})$  followed by a stretch by a factor 2. To see what exactly the condition is, one cranks out the algebra:

$$\begin{aligned} \mathbf{PP} &= \mathbf{xx}^T\mathbf{xx}^T + \mathbf{xx}^T\mathbf{yy}^T + \mathbf{yy}^T\mathbf{xx}^T + \mathbf{yy}^T\mathbf{yy}^T \\ &= \mathbf{xx}^T + \mathbf{x}(\mathbf{x}^T\mathbf{y})\mathbf{y}^T + \mathbf{y}(\mathbf{y}^T\mathbf{x})\mathbf{x}^T + \mathbf{yy}^T \\ &= \mathbf{P} + \mathbf{x}(\mathbf{x}^T\mathbf{y})\mathbf{y}^T + \mathbf{y}(\mathbf{y}^T\mathbf{x})\mathbf{x}^T \\ &= \mathbf{P} + \langle \mathbf{x}, \mathbf{y} \rangle (\mathbf{xy}^T + \mathbf{yx}^T) \end{aligned}$$

We see that a sufficient condition for  $\mathbf{PP} = \mathbf{P}$  is  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ , that is,  $\mathbf{x}$  and  $\mathbf{y}$  are orthogonal. Conversely, if  $\mathbf{PP} = \mathbf{P}$ , then  $\langle \mathbf{x}, \mathbf{y} \rangle (\mathbf{xy}^T + \mathbf{yx}^T) = \mathbf{0}$ . For this to be the case,  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$  or  $\mathbf{xy}^T + \mathbf{yx}^T = \mathbf{0}$ . The former means  $\mathbf{x}$  and  $\mathbf{y}$  are orthogonal. The latter implies  $\mathbf{xy}^T = -\mathbf{yx}^T$ , which means that the diagonal elements  $x_i y_i = -x_i y_i$ , hence  $x_i y_i = 0$ . This, in turn, implies once again that  $\mathbf{x}$  and  $\mathbf{y}$  are orthogonal:  $\sum x_i y_i = 0$ .

8. If  $\mathbf{P}$  ( $n \times n$ ) is idempotent and symmetric, does the same hold for  $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ ?

**A:** Yes, it does.  $(\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P}) = \mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P}\mathbf{P} = \mathbf{I} - \mathbf{P}$ .  $(\mathbf{I} - \mathbf{P})^T = \mathbf{I}^T - \mathbf{P}^T = \mathbf{I} - \mathbf{P}$ . Hence  $\mathbf{I} - \mathbf{P}$ , too, is an orthogonal projection. (It projects onto the orthogonal complement.)

9. If  $\mathbf{P}$  is idempotent, when is it non-singular?

**A:** Note that any “residual vector”  $\mathbf{y} - \mathbf{P}\mathbf{y}$  is in the null space of  $\mathbf{P}$ :  $\mathbf{P}(\mathbf{y} - \mathbf{P}\mathbf{y}) = \mathbf{P}\mathbf{y} - \mathbf{P}\mathbf{P}\mathbf{y} = \mathbf{0}$ . For  $\mathbf{P}$  to be non-singular, the null space of  $\mathbf{P}$  must consist of only the zero vector  $\mathbf{0}$ . Hence it is necessary that all residual vectors are zero:  $\mathbf{y} = \mathbf{P}\mathbf{y}$  for all  $\mathbf{y}$ . In other words,  $\mathbf{P} = \mathbf{I}$ . The identity is the only non-singular projection (idempotent linear mapping).

Here is another proof: If  $\mathbf{P}$  is idempotent and non-singular, let  $\mathbf{Q}$  be its inverse. We have

$$\mathbf{P}\mathbf{Q} = \mathbf{I} \Rightarrow \mathbf{P}\mathbf{P}\mathbf{Q} = \mathbf{I} \Rightarrow \mathbf{P}(\mathbf{P}\mathbf{Q}) = \mathbf{I} \Rightarrow \mathbf{P}\mathbf{I} = \mathbf{I} \Rightarrow \mathbf{P} = \mathbf{I}.$$

10. Can you give idempotence an intuitive meaning? It may help to look at the range space and null space of an idempotent map  $\mathbf{P}$ .

**A:** (A note on terminology: range space = image space. Also used is the term “column space”, but this is objectionable because when we talk about linear mappings, we do so without a particular basis and coordinate system in mind, whereas a column of a matrix contains the coordinates of the image of a basis vector under the linear map, hence refers to coordinates in a particular basis.)

Intuition behind idempotence:

(1) The range space  $\mathcal{R} = \{\mathbf{P}\mathbf{z} \mid \mathbf{z}\}$  is the space of fixed points under  $\mathbf{P}$ : If  $\mathbf{y} = \mathbf{P}\mathbf{z}$ , then  $\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{P}\mathbf{z} = \mathbf{P}\mathbf{z} = \mathbf{y}$ . Therefore, once a point is mapped to  $\mathcal{R}$ , it stays put.

(2) The null space  $\mathcal{N} = \{\mathbf{y} \mid \mathbf{P}\mathbf{y} = \mathbf{0}\}$  is exactly the space of residual vectors: If  $\mathbf{y} = \mathbf{z} - \mathbf{P}\mathbf{z}$ , then  $\mathbf{P}\mathbf{y} = \mathbf{P}(\mathbf{z} - \mathbf{P}\mathbf{z}) = \mathbf{P}\mathbf{z} - \mathbf{P}\mathbf{z} = \mathbf{0}$  and hence  $\mathbf{y} \in \mathcal{N}$ ; conversely, if  $\mathbf{P}\mathbf{y} = \mathbf{0}$ , then  $\mathbf{y} = \mathbf{y} - \mathbf{P}\mathbf{y} \in \mathcal{N}$ . The residual vectors represent the “projection direction”; they are what disappears after projection:  $\mathbf{y} = \mathbf{P}\mathbf{y} + (\mathbf{y} - \mathbf{P}\mathbf{y})$ , where  $\mathbf{P}\mathbf{y}$  stays put and  $(\mathbf{y} - \mathbf{P}\mathbf{y})$  disappears. This is not so for general linear maps that don’t have the property of idempotence.

In summary, idempotence is the projection property that says that points get dropped onto a projection space where they stay put.

(Note that we are not talking about “orthogonal projection”! In fact, a map that is

idempotent can project in “oblique directions.” See a question below that goes into the connection between orthogonality and symmetry of a projection — it’s the same.)

11. Assuming  $\mathbf{P}$  and  $\mathbf{A}$  are of the same size  $n \times n$  and  $\mathbf{A}$  is non-singular, what is the meaning of  $\mathbf{Q} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}$ ? (Hint: coordinate transformation.)

**A:** This type of manipulation describes a basis change and ensuing change of coordinates. The columns of the matrix  $\mathbf{A}$  contain the coordinates of the old basis vectors in terms of the new basis vectors: If  $\mathbf{e}_j^{old} = \sum_i A_{ij} \mathbf{e}_i^{new}$ , then the same vector  $\mathbf{y}$  has coordinates  $y_i^{new}$  and  $y_j^{old}$  in the respective bases:  $\mathbf{y} = \sum_i y_i^{new} \mathbf{e}_i^{new} = \sum_j y_j^{old} \mathbf{e}_j^{old}$ . Using  $\mathbf{A}$ , we get

$$\mathbf{y} = \sum_i y_i^{new} \mathbf{e}_i^{new} = \sum_j y_j^{old} \mathbf{e}_j^{old} = \sum_{ij} y_j^{old} A_{ij} \mathbf{e}_i^{new}$$

Writing  $\mathbf{y}^{new}$  and  $\mathbf{y}^{old}$  for the coordinate  $n$ -tuples for  $\mathbf{y}$  (the abstract vector) with regard to the respective bases, we now see that  $\mathbf{y}^{new} = \mathbf{A}\mathbf{y}^{old}$ . That is, even though the columns of  $\mathbf{A}$  represent the old basis vectors in the new basis (the basis transformation),  $\mathbf{A}$  also represents the new coordinates in terms of the old coordinates (the resulting coordinate transformation).

With this preparation (which should have been familiar from a basic linear algebra course), we can attack the problem of how matrices (the “coordinates” of linear maps in a given basis and coordinate system) transform when the basis is changed. The same linear map ‘ $P$ ’ is described in the old coordinate system by  $\mathbf{z}^{old} = \mathbf{P}^{old} \mathbf{y}^{old}$ , and in the new coordinate system by  $\mathbf{z}^{new} = \mathbf{P}^{new} \mathbf{y}^{new}$ . The problem is to find  $\mathbf{P}^{new}$  from  $\mathbf{P}^{old}$ . We know that  $\mathbf{y}^{new} = \mathbf{A}\mathbf{y}^{old}$  and  $\mathbf{z}^{new} = \mathbf{A}\mathbf{z}^{old}$ , hence

$$\begin{aligned} \mathbf{z}^{new} &= \mathbf{P}^{new} \mathbf{y}^{new} = \mathbf{P}^{new} \mathbf{A}\mathbf{y}^{old} \\ &= \mathbf{A}\mathbf{z}^{old} = \mathbf{A}\mathbf{P}^{old} \mathbf{y}^{old} \end{aligned}$$

Hence  $\mathbf{P}^{new} \mathbf{A} = \mathbf{A}\mathbf{P}^{old}$ , or  $\mathbf{P}^{new} = \mathbf{A}\mathbf{P}^{old} \mathbf{A}^{-1}$ . That is,  $\mathbf{Q} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}$  is the same linear map as  $\mathbf{P}$  but expressed in a new basis with coordinate transformation given by  $\mathbf{A}$ .

Obviously  $\mathbf{P}^{new} = \mathbf{A}\mathbf{P}^{old} \mathbf{A}^{-1}$  holds for arbitrary linear maps  $\mathbb{R}^n \rightarrow \mathbb{R}^n$ , not just projections.

12. If  $\mathbf{P}$  is idempotent, when is  $\mathbf{Q} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}$ , too?

**A:** Always.  $\mathbf{Q}\mathbf{Q} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}\mathbf{A}\mathbf{P}\mathbf{A}^{-1} = \mathbf{A}\mathbf{P}\mathbf{P}\mathbf{A}^{-1} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1} = \mathbf{Q}$ .

13. If  $\mathbf{P}$  is symmetric, when is  $\mathbf{Q} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}$ , too?

**A:** Not always!  $\mathbf{Q}^T = (\mathbf{A}^{-1})^T \mathbf{P}^T \mathbf{A}^T$  which equals  $\mathbf{Q} = \mathbf{A}\mathbf{P}\mathbf{A}^{-1}$  if  $\mathbf{A}^{-1} = \mathbf{A}^T$ , that is, if  $\mathbf{A}$  is an orthogonal matrix, that is, if the coordinate transformation is between orthonormal bases.

14. If the idempotent  $\mathbf{P}$  is symmetric, show that its range space and null space are orthogonal to each other. If you are ambitious, show also the converse (not mandatory).

**A:** We have to show orthogonality of the range space and null space, i.e.,  $\langle \mathbf{P}\mathbf{y}, \mathbf{z} - \mathbf{P}\mathbf{z} \rangle = 0$   $\forall \mathbf{y}, \mathbf{z}$  because all elements of the null space are of the form  $\mathbf{z} - \mathbf{P}\mathbf{z}$ . This identity follows immediately from idempotence and symmetry:  $\langle \mathbf{P}\mathbf{y}, \mathbf{z} - \mathbf{P}\mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{P}^T \mathbf{z} - \mathbf{P}^T \mathbf{P}\mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{P}\mathbf{z} - \mathbf{P}\mathbf{P}\mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{P}\mathbf{z} - \mathbf{P}\mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{0} \rangle = 0$ .

The converse is a little trickier: Start with the tautology  $\langle \mathbf{P}\mathbf{y}, \mathbf{P}\mathbf{z} \rangle = \langle \mathbf{P}\mathbf{y}, \mathbf{P}\mathbf{z} \rangle$  and add zero to both sides: add  $\langle \mathbf{P}\mathbf{y}, \mathbf{z} - \mathbf{P}\mathbf{z} \rangle$  to the left and  $\langle \mathbf{y} - \mathbf{P}\mathbf{y}, \mathbf{P}\mathbf{z} \rangle$  to the right. (Both terms are zero because  $\mathbf{y} - \mathbf{P}\mathbf{y}$  and  $\mathbf{z} - \mathbf{P}\mathbf{z}$  are in the null space of  $\mathbf{P}$ , and  $\mathbf{P}\mathbf{y}$  and  $\mathbf{P}\mathbf{z}$  are in the range space of  $\mathbf{P}$ , and we assume the two spaces to be orthogonal.) Contracting the inner products with the reverse distributive law yields  $\langle \mathbf{P}\mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{P}\mathbf{z} \rangle$ , that is, symmetry of  $\mathbf{P}$ .

15. If  $\mathbf{X}$  is of size  $n \times p$  ( $n \geq p$ ) and of rank  $p$ , let  $\mathbf{P} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ . Is  $\mathbf{P}$  idempotent and symmetric?

**A:** Yes, both. Straightforward matrix algebra using associativity of matrix multiplication and cancellations of matrices with their inverses.

16. What are the range space and the null space of  $\mathbf{P}$  in the previous problem?

**A:**  $\mathcal{R}(\mathbf{P}) = \mathcal{R}(\mathbf{X})$ ; the “ $\subset$ ” part is immediate from the definition of  $\mathbf{P}$ . Equality follows from  $\mathbf{P}\mathbf{X} = \mathbf{X}$ .  $\mathcal{N}(\mathbf{P})$  is the residual space of the regression onto  $\mathbf{X}$ .

17. If  $\mathbf{X}$  is of size  $n \times p$  ( $n \geq p$ ) and full rank, exactly when is  $\mathbf{P} = \mathbf{X}\mathbf{X}^T$  idempotent? When is this matrix symmetric?

**A:**  $\mathbf{P}$  is idempotent when  $\mathbf{X}\mathbf{X}^T \mathbf{X}\mathbf{X}^T = \mathbf{X}\mathbf{X}^T$ . This certainly holds if  $\mathbf{X}^T \mathbf{X} = \mathbf{I}$ , that is, if the columns of  $\mathbf{X}$  are orthonormal. For the converse one should really assume that  $\mathbf{X}$  is full rank  $p$ , in which case  $\mathbf{X}$  must have o.n. columns. —  $\mathbf{P}$  is always symmetric, hence if it is idempotent, it is automatically an orthogonal projection.

18. If  $\mathbf{P}$  is idempotent, when does the inverse of  $\mathbf{I} + c\mathbf{P}$  exist? What is it?

**A:**  $\mathbf{I} + c\mathbf{P}$  has no inverse when it is singular, that is, when there exists  $\mathbf{y} \neq \mathbf{0}$  such that  $\mathbf{y} + c\mathbf{P}\mathbf{y} = \mathbf{0}$ , i.e.,  $\mathbf{y} = -c\mathbf{P}\mathbf{y}$ . Applying  $\mathbf{P}$  to both sides and using idempotence we get  $\mathbf{P}\mathbf{y} = -c\mathbf{P}\mathbf{y}$ . Together with the previous equation we see that  $\mathbf{P}\mathbf{y} = \mathbf{y} \neq \mathbf{0}$ , hence from the last equation  $c = -1$ . — Thus for  $c \neq -1$  the linear map  $\mathbf{I} + c\mathbf{P}$  is non-singular and has an inverse. The inverse can be written down explicitly, and one intuition for finding it is to note that  $\mathbf{I} + c\mathbf{P}$  acts a little like  $\mathbf{P}$  in that it works one way for  $\mathbf{y} \in \mathcal{R}(\mathbf{P})$  and another way for  $\mathbf{y} \in \mathcal{N}(\mathbf{P})$ . One may therefore figure that the inverse is also of the form  $\mathbf{I} + d\mathbf{P}$ . If correct, what  $d$  is can be found out by checking the condition of being an inverse:  $\mathbf{I} = (\mathbf{I} + c\mathbf{P})(\mathbf{I} + d\mathbf{P}) = \mathbf{I} + (c + d)\mathbf{P} + cd\mathbf{P}\mathbf{P} = \mathbf{I} + (c + d + cd)\mathbf{P}$ , hence  $c + d + cd = 0$ , or  $d = -c/(1 + c)$ . The solution is:

$$(\mathbf{I} + c\mathbf{P})^{-1} = \mathbf{I} - \frac{c}{1 + c}\mathbf{P}$$

We see again that this inverse exists exactly for  $c \neq -1$ .

19. Subject the identity for the inverse of  $\mathbf{I} + c\mathbf{P}$  obtained in the previous item to a coordinate transformation with a matrix  $\mathbf{A}$  (as earlier). What general identity do you obtain?

**A:**

$$\begin{aligned} \mathbf{A}(\mathbf{I} + c\mathbf{P})^{-1}\mathbf{A}^{-1} &= \mathbf{A}\left(\mathbf{I} - \frac{c}{1 + c}\mathbf{P}\right)\mathbf{A}^{-1} \\ (\mathbf{I} + c\mathbf{A}\mathbf{P}\mathbf{A}^{-1})^{-1} &= \mathbf{I} - \frac{c}{1 + c}\mathbf{A}\mathbf{P}\mathbf{A}^{-1} \end{aligned}$$

This is nothing new, though, because  $\mathbf{A}\mathbf{P}\mathbf{A}^{-1}$  is just the same idempotent map (oblique projection) expressed in different coordinates.