Reminder/summary/hint of projection problem

I will summarize here different approaches to the projection problem. If the problem is on your mind, and you want to go further with it now, please go ahead. But otherwise I would recommend skipping this section for now and coming back to it when you have time. Note that nothing after this section depends logically on the projection problem or least squares (and we will come back to them).

1. (Core (but could do later)) I gave hints to a different approach to the projection problem in Class #6, Friday March 5 (see notes). I'll quickly summarize the hint here: if we are projecting a vector $\vec{v}$ in vector space $V = \mathbb{R}^3$ perpendicularly down to a 2-dimensional subspace (i. e. plane through origin) V', then the projection $P\vec{v}$ has to satisfy two conditions:
   (i) $P\vec{v} = A\vec{t}$
   (ii) $A^\mathsf{T}(\vec{v} - P\vec{v}) = \vec{0}$
   where A is a matrix whose columns are the vectors $\vec{u}_1$, $\vec{u}_2$ spanning the subspace V', and $\vec{t} \in \mathbb{R}^2$. The first condition expresses that $P\vec{v}$ must lie in the subspace V', and the second condition expresses that $\vec{v} - P\vec{v}$ must be orthogonal to the subspace V'.
   (a) Use these hints to write a solution to the projection problem in matrix form. That is, find $P\vec{v}$, in general form, as a matrix multiplication. (You may find it helpful to write out the entries of the matrices, to understand what's going on; but the final answer will look simpler if you just write it in matrices abstractly, as I did in conditions (i) and (ii) above.)
   (b) The projection operator $P : V \to V$ can in fact be written as a matrix. Use your answer to the previous part to write the matrix for P in terms of A.
   (c) Check the dimensions of matrices in your formula to make sure they make sense. (What is the dimension of the input vector and output vector for each matrix multiplication?)
   (d) (Challenge) For the solution to work, we need to know that the matrix $A^\mathsf{T}A$ is invertible. (Why?) How do we know that this is the case? Or, if it isn't always the case, what is the condition?
   (e) Try it out numerically. (If you have been working on a numerical example previously, try computing it here.) Check that your answer obeys the conditions on the projection!
   (f) Does this solution still work if $V = \mathbb{R}^n$, (with the standard dot product), and V' is some k-dimensional subspace spanned by linearly independent vectors $\vec{u}_1, \ldots, \vec{u}_k$? What are the corresponding dimensions of the matrices—do all the matrix multiplications make sense?

Fields, Vector Spaces, Bases, Linear Transformations

2. (Core) (A simple example of all concepts so far) I want to go through an example which will illustrate most of the abstract concepts introduced so far. I would recommend doing this problem with some care: I think it will help keep you oriented with the more abstract concepts and theorems.

Suppose that $V = \mathbb{R}^2$ is our vector space. Let $\vec{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix}1\\0\end{bmatrix}$. Then $B = \{\vec{e}_1, \vec{e}_2\}$ is called the *standard basis* of $\mathbb{R}^2$. [1]

Let $\vec{e}'_1 = \begin{bmatrix}1\\0\end{bmatrix}$, $\vec{e}'_2 = \begin{bmatrix}1\\1\end{bmatrix}$. Then $B' = \{\vec{e}'_1, \vec{e}'_2\}$ is also a basis of $\mathbb{R}^2$ (which we will prove shortly).

Suppose we are given some vector $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix}$, where x and y are some given real numbers. If we can write $\vec{v}$ as a linear combination of the elements of B':

$$\vec{v} = x'\vec{e}'_1 + y'\vec{e}'_2,$$

or

$$\begin{bmatrix}x\\y\end{bmatrix} = x' \begin{bmatrix}1\\0\end{bmatrix} + y' \begin{bmatrix}1\\1\end{bmatrix}, \tag{1}$$

where x' and y' are some real numbers, then x' and y' are called the *components* of $\vec{v}$ in the B' basis. We also define, for any x' and y', the notation

$$\begin{bmatrix}x'\\y'\end{bmatrix}_{B'} := x'\vec{e}'_1 + y'\vec{e}'_2,$$

and similarly for B; so if x', y', x, y satisfy equation (1), then we have

$$\vec{v} = \begin{bmatrix}x'\\y'\end{bmatrix}_{B'} = \begin{bmatrix}x\\y\end{bmatrix}_{B}.$$

This is the *same* vector, expressed in two *different* bases (or coordinate systems).

In $\mathbb{R}^2$ (or more generally k^n), we will often drop the subscript for the standard basis, so

$$\vec{v} = \begin{bmatrix}x'\\y'\end{bmatrix}_{B'} = \begin{bmatrix}x\\y\end{bmatrix}_{B} = \begin{bmatrix}x\\y\end{bmatrix}.$$

(but that option only exists in k^n, not in other vector spaces).

For this example, I will define subspaces

$$U = \operatorname{span}\vec{e}_1 = \operatorname{span}\vec{e}'_1$$

and

$$W = \operatorname{span}\vec{e}_w, \quad W' = \operatorname{span}\vec{e}'_2;$$

then $V = U \oplus W = U \oplus W'$ (which we will prove shortly). I will also write P_U^W for the projection to subspace U along complementary subspace W (see the notes for the definition).

(a) Draw a grid for the standard basis B, and the corresponding (x, y) coordinate axes. (That is, draw in the grid lines for the coordinate system.) Now, in a different color, but overlaid on the *same* axes, draw in the grid lines and axes for the B' basis and (x', y') coordinate system. It doesn't have to be perfect, but do this relatively carefully.

[1] We can similarly define a *standard basis* for any vector space defined as $V = k^n$ for some field k and positive integer n. But more general vector spaces don't have any particular standard basis. I've called it a standard "basis", but I haven't proved it is a basis; it's not difficult to prove, and I'll ask you to do that later.

(b) In a third color, draw the vector $\vec{v} = \begin{bmatrix}3\\2\end{bmatrix}_{B'}$, i. e. the vector with endpoint coordinates $x' = 3$, $y' = 2$. Use your diagram to find the standard B components (i. e. standard x and y coordinates) of this vector. Verify that your answer is correct using equation (1). For this vector, what is $P_U^{W'}\vec{v}$? (Remember that, by definition, the projection is a *vector*, not a scalar; the corresponding *coordinate* is a scalar.) What is $P_U^W\vec{v}$? Note that $P_U^W\vec{v} \neq P_U^{W'}\vec{v}$; the projection to U depends on the choice of complementary subspace. What are $P_{W'}^U\vec{v}$ and $P_W^U\vec{v}$?

(c) On the same grid(s), draw the vector $\vec{v} = \begin{bmatrix}1\\3\end{bmatrix}_B$. Do all the similar steps to the previous part: use the graph to find the expression in basis B', (i. e. find the x' and y' coordinates), check with equation (1), and find all the corresponding projections.

(d) Write formulas for the x and y coordinates in terms of the x' and y' coordinates. (This is just a restatement of (1).)

(e) If $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix}_B$ is given, write a formula for its x' and y' coordinates. Also write a formula for the vectors $P_U^{W'}\vec{v}$ and $P_{W'}^U\vec{v}$ (your answer will be vectors involving the givens x and y).

(f) In the previous part, you found $P_U^{W'}\vec{v}$ and $P_{W'}^U\vec{v}$ for a given vector $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix}_B$. Write each of these two formulas in matrix form, that is, write $P_U^{W'}\vec{v}$ as a matrix with known coefficients, times the vector $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix}_B$, and similarly for $P_{W'}^U\vec{v}$. (What should the dimensions of the matrices be?) (I'm getting ahead of myself a bit here. We will formally introduce matrices in the next problem set. But I will assume for the moment that you remember enough about matrices to answer this question.)

(g) What do you get if you add the matrices

$$P_U^{W'}\vec{v} + P_{W'}^U\vec{v}\ ?$$

What does this answer mean?

(h) Given a vector $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix}_B$, write your formulas for the coordinates x' and y' in matrix form. Since the answers for x' and y' are scalars, the formulas should be in the form of a known 1×2 matrix times the column vector $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix}_B$.

(i) Put your answers for the previous part back into equation (1), and show how this gives another way of finding the projection matrices $P_U^{W'}$ and $P_{W'}^U$. (It's a little tricky: I claim if you set it up correctly, these matrices will be the product of a 2×1 matrix, times a 1×2 matrix.)

(j) Answer all the questions, starting from question (**??**), again for the projections P_U^W and P_W^U. These are the projections to the standard basis vectors; give all the answers in terms of standard coordinates. (The answers should look pretty simple.)

3. (CORE) Let $V = \mathbb{R}_3[x]$, where

$$\mathbb{R}_3[x] = \{a_0 + a_1x + a_2x^2 + a_3x^3 \Big| a_0, a_1, a_2 \in \mathbb{R}\}$$

is the vector space of real polynomials of degree less than or equal to 3. The first four Legendre polynomials form a basis for this space; they are

$$P_0(x) = 1,\ P_1(x) = x,\ P_2(x) = \frac{3}{2}x^2 - \frac{1}{2},\ P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x.$$

(There are in fact infinitely many of them, which form a basis for the vector space $\mathbb{R}[x]$ of all real polynomials; these occur in the quantum mechanics of finding the shape of electron shells in atoms, among many other places.) Write the polynomial $1 - 2x + 3x^2 + x^3$ as a linear combination of the elements of $B = \{P_0, \ldots, P_4\}$.

4. (OPTIONAL, but interesting; both theory and computation) This question is about *Lagrange interpolation*, which is an efficient way of exactly fitting a polynomial to some given data points.
   (a) Suppose that we have two data points (x_1, y_1) and (x_2, y_2), $x_1 \neq x_2$, and we want to fit a line to them. I want to imagine that the x_1 and x_2 are fixed, but that the y_1 and y_2 might vary, and that I want to quickly find the unique straight line that goes through (x_1, y_1) and (x_2, y_2) for a variety of different y_1 and y_2, keeping x_1 and x_2 fixed. Define the *Lagrange polynomials* ℓ_1 and ℓ_2 for this situation by

$$\ell_1(x) = \frac{x - x_2}{x_1 - x_2},\ \ell_2(x) = \frac{x - x_1}{x_2 - x_1}$$

   I claim that these two polynomials form a basis for the vector space $\mathbb{R}_1[x]$ of all straight line functions. Show that the linear combination

$$y_1\ell_1 + y_2\ell_2$$

   gives a polynomial (straight line) which passes through the given data points (x_1, y_1) and (x_2, y_2).
   (b) Suppose now that we have three data points, (x_1, y_1), (x_2, y_2), and (x_3, y_3), with none of the x_i equal to each other. In this case, we define the Lagrange polynomials to be

$$\ell_1(x) = \frac{x - x_2}{x_1 - x_2}\frac{x - x_3}{x_1 - x_3},\ \ell_2(x) = \frac{x - x_1}{x_2 - x_1}\frac{x - x_3}{x_2 - x_3},\ \text{and}\ \ell_2(x) = \frac{x - x_1}{x_3 - x_1}\frac{x - x_2}{x_3 - x_2}.$$

   I claim that these three polynomials form a basis for the vector space $\mathbb{R}_2[x]$ of all quadratic functions. Show that the linear combination

$$y_1\ell_1 + y_2\ell_2 + y_3\ell_3$$

   gives a polynomial (parabola) which passes through the given data points (x_1, y_1), (x_2, y_2), and (x_3, y_3).

5. (OPTION: THEORY, but everyone should understand the statement) Suppose that V is a vector space, with subspaces V' and V''. Prove that $V = V' \oplus V''$ (see the notes for definition) if, and only if, both
   (i) $V = V' + V''$, and
   (ii) $V' \cap V'' = \{\vec{0}\}$.

6. (OPTION:THEORY, but everyone should understand the statement) Suppose that V is a vector space, with subspaces V' and V'', such that $V = V' \oplus V''$. Prove that the projection $P_{V'}^{V''}$, to the subspace V' along the subspace V'', is a linear transformation
$$P_{V'}^{V''} : V \to V.$$
Also, prove that
$$P_{V'}^{V''} + P_{V''}^{V'} = I,$$
where $I : V \to V$ is the identity transformation, $I\vec{v} = \vec{v}$ for all $\vec{v} \in V$.
7. (OPTIONAL, both theory and computation) This is a question about finite fields, which are important both in math and computer science. Suppose that k is a finite field; that is, k is a field, and as a set, it has finitely many elements. Let's say k has m elements.
   (a) How many elements does the vector space k^n have?
   (b) Prove that if k_0 is a subfield of k, then k is a vector space over k_0. (This is true for any field k, not just finite ones.)
   (c) (Challenge) Prove that any finite field has a subfield of order some prime p. (Hint: consider the set of all elements of the form $1 + 1 + \cdots + 1$.)
   (d) Prove that the order of a finite field k must be p^n, for some prime number p and positive integer n.