1. Let \( \mathbf{x}_1 = j_4, \mathbf{x}_2 = (4, 1, 3, 4)^T, \mathbf{y} = (1, 9, 5, 5)^T. \) Let \( V = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2). \)

   a. Find \( \hat{\mathbf{y}} = p(\mathbf{y}|V) \) and \( \mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}. \)
   b. Find \( \hat{\mathbf{y}}_1 = p(\mathbf{y}|\mathbf{x}_1) \) and \( \hat{\mathbf{y}}_2 = p(\mathbf{y}|\mathbf{x}_2) \) and show that \( \hat{\mathbf{y}} \neq \hat{\mathbf{y}}_1 + \hat{\mathbf{y}}_2. \)
   c. Verify that \( \mathbf{e} \perp V. \)
   d. Find \( ||\mathbf{y}||^2, ||\hat{\mathbf{y}}||^2, ||\mathbf{e}||^2, \) and verify that the Pythagorean Theorem holds. Compute \( ||\hat{\mathbf{y}}||^2 \) directly from \( \hat{\mathbf{y}} \) and also by using the formula \( ||\hat{\mathbf{y}}||^2 = \mathbf{y}^T \mathbf{P} \mathbf{y} \)
   where \( \mathbf{P} \) is the projection matrix onto \( V. \)
   e. Use Gram-Schmidt orthogonalization to find four mutually orthogonal vectors \( \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \) and \( \mathbf{v}_4 \) such that \( V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2). \) Hint: You can choose \( \mathbf{x}_3 \) and \( \mathbf{x}_4 \) arbitrarily, as long as \( \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \) are LIN.

2. (Simple linear regression.) Let \( \mathbf{y} = (y_1, \ldots, y_n)^T, \mathbf{x} = (x_1, \ldots, x_n)^T, \) and \( V = \mathcal{L}(\mathbf{j}_n, \mathbf{x}). \)

   a. Use Gram-Schmidt orthogonalization on the vectors \( \mathbf{j}_n, \mathbf{x} \) (in this order) to find orthogonal vectors \( \mathbf{j}_n, \mathbf{x}^* \) spanning \( V. \) Express \( \mathbf{x}^* \) in terms of \( \mathbf{j}_n \) and \( \mathbf{x} \), then find \( b_0, b_1 \) such that \( \hat{\mathbf{y}} = b_0 \mathbf{j}_n + b_1 \mathbf{x}. \) To simplify the notation, let \( \mathbf{y}^* = \mathbf{y} - p(\mathbf{y}|\mathbf{j}_n) = \mathbf{y} - \hat{\mathbf{y}}_n, \)

   \[
   S_{xy} = \langle \mathbf{x}^*, \mathbf{y}^* \rangle = \sum_i (x_i - \bar{x})(y_i - \bar{y}) = \sum_i (x_i - \bar{x})y_i = \sum_i x_i y_i - n \bar{x} \bar{y},
   \]

   \[
   S_{xx} = \langle \mathbf{x}^*, \mathbf{x}^* \rangle = \sum_i (x_i - \bar{x})^2 = \sum_i (x_i - \bar{x})x_i = \sum_i x_i^2 - n \bar{x}^2,
   \]

   \[
   S_{yy} = \langle \mathbf{y}^*, \mathbf{y}^* \rangle = \sum_i (y_i - \bar{y})^2.
   \]

   b. Suppose \( \hat{\mathbf{y}} = p(\mathbf{y}|V) = a_0 \mathbf{j}_n + a_1 \mathbf{x}^*. \) Find formulas for \( a_i \) and \( A_0 \) in terms of \( \hat{\mathbf{y}}, S_{xy}, \) and \( S_{xx}. \)

   c. Express \( \mathbf{x}^* \) in terms of \( \mathbf{j}_n \) and \( \mathbf{x} \), and use this to determine formulas for \( b_0 \) and \( b_1 \) so that \( \hat{\mathbf{y}} = b_0 \mathbf{j}_n + b_1 \mathbf{x}. \)

   d. Express \( ||\hat{\mathbf{y}}||^2 \) and \( ||\mathbf{y} - \hat{\mathbf{y}}||^2 \) in terms of \( S_{xy}, S_{xx} \) and \( S_{yy}. \)

   e. Use the formula \( \mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \) for \( \mathbf{b} = (b_0, b_1)^T, \) and verify that this gives the same answer as in (c).

   f. for \( \mathbf{y} = (2, 6, 8, 8)^T, \mathbf{x} = (0, 1, 2, 3)^T \) find \( a_0, a_1, \hat{\mathbf{y}}, b_0, b_1, ||\mathbf{y}||^2, ||\hat{\mathbf{y}}||^2, ||\mathbf{y} - \hat{\mathbf{y}}||^2. \) Verify that \( ||\hat{\mathbf{y}}||^2 = b_0 \langle \mathbf{y}, \mathbf{j}_4 \rangle + b_1 \langle \mathbf{y}, \mathbf{x} \rangle \) and that \( \langle \mathbf{y} - \hat{\mathbf{y}}, V \rangle = 0. \)

3. Let \( x_1, \ldots, x_k \) be a basis of a subspace \( V \subset \mathbb{R}^n. \) Suppose that \( p(\mathbf{y}|V) = \sum_{j=1}^k p(\mathbf{y}|x_j) \) for every vector \( \mathbf{y} \in \mathbb{R}^n. \) Prove that \( x_1, \ldots, x_k \) are mutually orthogonal. Hint: Consider the vector \( \mathbf{y} = x_i \) for each \( i. \)

4. Show that for \( \mathbf{X}_{n \times k} = \mathbf{X}_{n \times k} \mathbf{B}_{k \times k} \) with \( \mathbf{B} \) nonsingular and \( \mathbf{X} \) of full rank, \( \mathbf{P} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \) remains unchanged if \( \mathbf{X} \) is replaced by \( \mathbf{W}. \) Thus \( \mathbf{P} \) is a function of the subspace spanned by the columns of \( \mathbf{X}, \) not of the particular basis chosen for this subspace (we can change \( \mathbf{X} \) without affecting \( \mathbf{P} \) as long as we haven't changed \( C(\mathbf{X}). \)

5. For each subspace \( V \) of \( \mathbb{R}^3 \) give the corresponding projection matrix \( \mathbf{P}. \) In each case verify that \( \mathbf{P} \) is symmetric and idempotent.

   a. \( V = \mathcal{L}(\mathbf{x}) \) where \( \mathbf{x} = (2, -1, -1)^T. \)
   b. \( V = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2) \) where \( \mathbf{x}_1 = (1, 1, 1)^T, \) and \( \mathbf{x}_2 = (1, -1, 0)^T. \)

6. For the subspace \( V = \mathcal{L}(\mathbf{j}_n, \mathbf{x}^*) \) of problem 2, what is \( \mathbf{P}_V? \) (Note that \( V = \mathcal{L}(\mathbf{j}_n, \mathbf{x}^*), \) also.) What is \( \mathbf{P}_{V^\perp}? \)