MT2.1 (37 Points) Consider a square $n \times n$ matrix $\mathbf{P}$ having *all* of the following properties:

- Each entry of $\mathbf{P}$ is either 0 or 1.
- Each column of $\mathbf{P}$ contains exactly one 1.
- Each row of $\mathbf{P}$ contains exactly one 1.

(a) (10 Points) Show that $\mathbf{P}^\mathsf{T}\,\mathbf{P} = \mathbf{I}$.

We'll accept any logically-sound explanation. You need not resort to elaborate mathematical derivations. It's possible to explain this result in two to three sentences.

Even if you have trouble showing this result, you may use it in the subsequent parts of the problem, if you find it useful.

(b) (13 Points) For this part only, consider the matrices

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} e & 0 & d \\ 0 & a & 0 \\ c & 0 & b \end{bmatrix}.$$

Determine the matrix $\mathbf{PAP}^\mathsf{T}$. Your final answer must be in the form of a single 3×3 matrix—not a product, sum, or other nontrivial form.

(c) (14 Points) Consider a vector $\boldsymbol{x}$ in $\mathbb{R}^n$ whose mean is

$$\mu_x = \frac{x_1 + \cdots + x_n}{n} = \frac{1}{n}\langle \mathbf{1}, \boldsymbol{x} \rangle = \frac{1}{n}\mathbf{1}^\mathsf{T}\,\boldsymbol{x},$$

and whose variance is

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)^2 = \frac{||\boldsymbol{x} - \mu_x \mathbf{1}||^2}{n} = \frac{||\boldsymbol{x}||^2}{n} - \mu_x^2.$$

Let $\boldsymbol{y} = \mathbf{P}\boldsymbol{x}$, and denote the mean and variance of the vector $\boldsymbol{y}$ by μ_y and σ_y^2, respectively.

(i) Show that $\mu_y = \mu_x$.

(ii) Determine a simple expression for σ_y^2, the variance of $\boldsymbol{y}$.
Your expression must be in closed form (not summations $\sum$), and depend *at most* on $\mathbf{P}$, μ_x, and σ_x^2.

(iii) Determine a simple expression for σ_{xy}, the covariance of $\boldsymbol{x}$ and $\boldsymbol{y}$, where the covariance is defined as

$$\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y) = \frac{\langle \boldsymbol{x} - \mu_x\mathbf{1}, \boldsymbol{y} - \mu_y\mathbf{1} \rangle}{n} = \frac{\boldsymbol{x}^\mathsf{T}\,\boldsymbol{y}}{n} - \mu_x\mu_y.$$

Your expression must be in closed form (not summations $\sum$), and depend *at most* on $\boldsymbol{x}$, $\mathbf{P}$, μ_x, and σ_x.

MT2.2 (37 Points) Consider a pair of real-valued 2-vectors $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, where $x_1 \neq x_2$ and $y_1 \neq y_2$. The respective means of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are given by

$$\mu_x = \mathsf{avg}(\boldsymbol{x}) = \frac{x_1 + x_2}{2} \qquad \mu_y = \mathsf{avg}(\boldsymbol{y}) = \frac{y_1 + y_2}{2}.$$

The respective de-meaned vectors of $\boldsymbol{x}$ and $\boldsymbol{y}$ are

$$\tilde{\boldsymbol{x}} = \boldsymbol{x} - \mu_x \mathbf{1} \qquad \text{and} \qquad \tilde{\boldsymbol{y}} = \boldsymbol{y} - \mu_y \mathbf{1}.$$

And the correlation coefficient of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is

$$\rho_{xy} \triangleq \frac{\langle \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \rangle}{||\tilde{\boldsymbol{x}}|| \, ||\tilde{\boldsymbol{y}}||} = \frac{\tilde{\boldsymbol{x}}^\top \tilde{\boldsymbol{y}}}{||\tilde{\boldsymbol{x}}|| \, ||\tilde{\boldsymbol{y}}||}.$$

The symbol $\triangleq$ signifies a definition.

(a) (13 Points) Show that $\rho_{xy} = \pm 1$—the correlation coefficient of $\boldsymbol{x}$ and $\boldsymbol{y}$ is either $+1$ or -1.

(b) (12 Points) For this part only, suppose $\boldsymbol{y} = \alpha\, \boldsymbol{x} + \beta \mathbf{1}$, where α and β are real parameters, and $\mathbf{1}$ is the all-ones vector in $\mathbb{R}^2$.

(i) In a short sentence, explain why α cannot be zero.

(ii) Explain why it must be that $\beta = \mu_y - \alpha\, \mu_x$, and then show

$$\rho_{xy} = \begin{cases} -1 & \text{if } \alpha < 0 \\ +1 & \text{if } \alpha > 0 \end{cases}.$$

(c) (12 Points) True or False? Any pair of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^2$, such that $x_1 \neq x_2$ and $y_1 \neq y_2$, can be expressed as affine functions of each other.

Provide a succinct, yet clear and convincing explanation.

MT2.3 (37 Points) Consider the following vectors in $\mathbb{R}^3$:

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

(a) (21 Points) Use the Gram-Schmidt Algorithm to determine a set of three vectors $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, and $\boldsymbol{q}_3$ such that

- For every $k, \ell = 1, 2, 3$, we have

$$\langle \boldsymbol{q}_k, \boldsymbol{q}_\ell \rangle = \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{if } k \neq \ell. \end{cases}$$

- For every $\ell = 1, 2, 3$, we have

$$\mathsf{span}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell) = \mathsf{span}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_\ell).$$

- Each $\boldsymbol{a}_\ell$ is a linear function of $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_\ell$ for all $\ell = 1, 2, 3$.
- Each $\boldsymbol{q}_\ell$ is a linear function of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell$ for all $\ell = 1, 2, 3$.

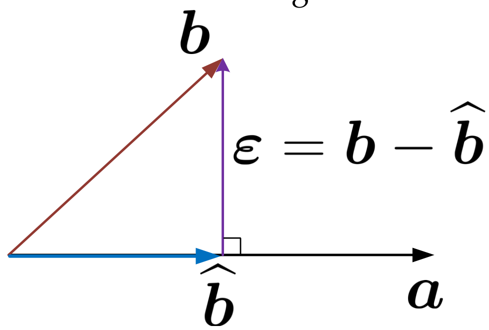
(b) (16 Points) Solve the equation $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ for the unknown vector $\boldsymbol{x}$, where

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{bmatrix} 2 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

You must explain the steps of your work.

MT2.4 (37 Points) Consider a nonzero vector $\boldsymbol{a}$ in $\mathbb{R}^n$, and another nonzero vector $\boldsymbol{b}$ that is not in the linear span of $\boldsymbol{a}$. We want to determine a vector $\widehat{\boldsymbol{b}}$—restricted to the linear span of $\boldsymbol{a}$—to serve as an estimate of $\boldsymbol{b}$.

To make $\widehat{\boldsymbol{b}}$ a reasonable estimate, we insist that it correspond to the point in $\mathsf{span}\,(\boldsymbol{a})$ closest to $\boldsymbol{b}$. If we denote the estimation error vector by $\boldsymbol{\varepsilon} = \boldsymbol{b} - \widehat{\boldsymbol{b}}$, Pythagoras's Theorem tells us that the error vector having the smallest Euclidean norm must be orthogonal to $\boldsymbol{a}$. This is illustrated in the diagram below.



Note that since $\widehat{\boldsymbol{b}} \in \mathsf{span}(a)$, we can represent it as $\widehat{\boldsymbol{b}} = \mu\boldsymbol{a}$ for some real scalar μ.

(a) (7 Points) The estimate $\widehat{\boldsymbol{b}}$ can be obtained from $\boldsymbol{b}$ using a linear transformation $\widehat{\boldsymbol{b}} = \mathbf{P}\boldsymbol{b}$, where $\mathbf{P}$ is a real-valued $n \times n$ matrix. Show that

$$\mathbf{P} = \boldsymbol{a}\left(\boldsymbol{a}^{\mathsf{T}}\boldsymbol{a}\right)^{-1}\boldsymbol{a}^{\mathsf{T}}.$$

To do this, enforce the orthogonality of $\boldsymbol{\varepsilon}$ and $\boldsymbol{a}$ to show that the value of μ that minimizes $||\boldsymbol{\varepsilon}||$ is given by

$$\mu = \left(\boldsymbol{a}^{\mathsf{T}}\boldsymbol{a}\right)^{-1}\boldsymbol{a}^{\mathsf{T}}\boldsymbol{b}.$$

Then determine $\mathbf{P}$.

(b) (5 Points) Determine $\mathsf{rank}(\mathbf{P})$, the rank of $\boldsymbol{P}$—that is, the number of independent columns (or independent rows)—of $\boldsymbol{P}$.

Explain how you arrive at your answer.

(c) (7 Points) This part explores some of the properties of the matrix $\mathbf{P}$.

(i) Show that $\mathbf{P}^{\mathsf{T}} = \mathbf{P}$.

(ii) Show that $\mathbf{P}^k = \mathbf{P}$, for all $k = 1, 2, 3, \ldots$.
In one or two sentences, provide a geometric interpretation of this result.

(d) (5 Points) Suppose a vector $\boldsymbol{y}$ is such that $\mathbf{P}\boldsymbol{y} = \mathbf{0}$. Determine

$$\langle \boldsymbol{a}, \boldsymbol{y} \rangle = \boldsymbol{a}^{\mathsf{T}}\boldsymbol{y}.$$

Explain the work toward your result, and then provide a one-sentence interpretation of what your result implies about how the vectors $\boldsymbol{a}$ and $\boldsymbol{y}$ are oriented relative to one another.

(e) (13 Points) Consider the matrix $\boldsymbol{S} = \mathbf{I} - \mathbf{P}$.

(i) Determine $\mathbf{S}^k$ for all $k = 2, 3, 4, \ldots$.
Hint: First determine $\mathbf{S}^2$. Higher powers will then become apparent.

(ii) Suppose a vector $\boldsymbol{y}$ is such that $\mathbf{P}\boldsymbol{y} = \mathbf{0}$. Determine the vector $\mathbf{S}\boldsymbol{y}$.
Provide a succinct, yet clear and convincing explanation of your work.

(iii) Show that $\{\boldsymbol{a}\}$ forms a basis for $\mathcal{N}(\mathbf{S})$, the nullspace of $\mathbf{S}$. The nullspace of $\mathbf{S}$ is the defined as the subspace

$$\mathcal{N}(\mathbf{S}) \triangleq \{\boldsymbol{v} \in \mathbb{R}^n \,|\, \mathbf{S}\boldsymbol{v} = \mathbf{0}\}.$$

The symbol $\triangleq$ signifies a definition.
Hint: Show that any vector $\boldsymbol{v}$ for which $\mathbf{S}\boldsymbol{v} = \mathbf{0}$ must be linearly dependent with $\boldsymbol{a}$.

(iv) Determine $\mathsf{rank}(\mathbf{S})$, where $\mathsf{rank}(\mathbf{S})$—called the rank of $\mathbf{S}$—is the number of independent columns (or independent rows)—of $\mathbf{S}$.
Provide a succinct, yet clear and convincing explanation of your work.

MT2.5 (37 Points)

Consider the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$.

Define $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}} \triangleq \boldsymbol{x}^{\mathsf{T}}\, \mathbf{W} \boldsymbol{y}$, where $\mathbf{W}$ is an $n \times n$ real-valued matrix. The symbol $\triangleq$ signifies a definition.

This problem explores how $\mathbf{W}$ determines whether $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ can serve as an inner product.

Recall that any inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ must satisfy the following properties:

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y} \rangle &= \langle \boldsymbol{y}, \boldsymbol{x} \rangle \\
\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle &= \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle \\
\langle c\, \boldsymbol{x}, \boldsymbol{y} \rangle &= c\, \langle \boldsymbol{x}, \boldsymbol{y} \rangle \\
\langle \boldsymbol{x}, \boldsymbol{x} \rangle &\geq 0, \;\; \text{with equality if, and only if, } \boldsymbol{x} = 0.
\end{aligned}$$

In one or more parts below, you may find it useful to study $\boldsymbol{z}^{\mathsf{T}}\, \mathbf{W} \boldsymbol{z}$, where

$$\boldsymbol{z} = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2.$$

(a) (7 Points) Determine an $n \times n$ matrix $\mathbf{W}$ such that $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ is the inner product that induces the familiar Euclidean norm on $\mathbb{R}^n$. In other words, determine $\mathbf{W}$ such that $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}} = \boldsymbol{x}^{\mathsf{T}}\, \boldsymbol{y}$.

(b) (7 Points) Show that if $\mathbf{W} = \begin{bmatrix} 0 & 3 \\ 1 & 0 \end{bmatrix}$, then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ is *not* an inner product on $\mathbb{R}^2$.

(c) (9 Points) Show that if $\mathbf{W} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ can serve as an inner product on $\mathbb{R}^2$.

Hint: It may help you to first show that $\langle \boldsymbol{z}, \boldsymbol{z} \rangle_{\mathbf{W}}$ can be expressed as

$$\langle \boldsymbol{z}, \boldsymbol{z} \rangle_{\mathbf{W}} = \lambda (a - b)^2 + \mu (a + b)^2$$

for appropriately-determined positive coefficients λ and μ.

(d) (7 Points) Show that if any of the diagonal entries of $\mathbf{W}$ is 0, then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ can*not* serve as an inner product on $\mathbb{R}^n$. In other words, if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ is a valid inner product, then the diagonal entries of $\mathbf{W}$ must be nonzero.

Hint: Assume, without loss of generality, that the entry w_{11} of the matrix $\mathbf{W}$ is zero. Then choose a *nonzero* vector $\boldsymbol{z}$ judiciously, such that $\boldsymbol{z}^{\mathsf{T}}\, \mathbf{W} \boldsymbol{z}$ violates at least one property of an inner product.

(e) (7 Points) Show that if $\mathbf{W}$ is *not* invertible, then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbf{W}}$ is *not* an inner product.