Problem 1 (20 pts) Suppose we are given three data points in $\mathbb{R}^2$, $(x,y) = (-\pi/2, 1), (0,1), (\pi/2, -1)$. Now, we want to find the best function to fit these points in the form of $y = \alpha + \beta \sin x$ in the sense of the least squares.

(a) (7 pts) Write a system of equation in the form $A\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ as if this function passes through all these three points.

Answer: Since $\sin x = -1, 0, 1$ when $x = -\pi/2, 0, \pi/2$, respectively, the system of equation can be easily written as:

$$\boxed{\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}}.$$

(b) (7 pts) Solve the least squares problem using the normal equation, and write the solution in the form of $y = \alpha + \beta \sin x$.

Answer: This is simply can be solved by

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (A^T A)^{-1} A^T \boldsymbol{b} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -1 \end{bmatrix}.$$

Thus the least square solution is:

$$\boxed{y = \frac{1}{3} - \sin x}.$$

(c) (6 pts) Let $\hat{\boldsymbol{x}}$ be the least squares solution you computed in (b). Compute the residual (or error) vector $\boldsymbol{b} - A\hat{\boldsymbol{x}}$ and its length in 2-norm. Compare this error size o the error size of the case $\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and confirm that the error of the least squares solution is smaller.

Answer: The error vector is simply,

$$\boldsymbol{b} - A\hat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/3 \\ -1 \end{bmatrix} = \boxed{\begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix}}.$$

Thus its 2-norm is:

$$\left\| \begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \end{bmatrix} \right\| = \boxed{\frac{\sqrt{6}}{3}},$$

which is in fact smaller than

$$\left\| \boldsymbol{b} - A \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \right\| = 3 > \frac{\sqrt{6}}{3}.$$

Problem 2 (20 pts) Consider the following matrix:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

(a) (5 pts) Find a basis of $\mathcal{R}(A)$. What is the rank of this matrix?

Answer: Let's compute E_A, the RREF of A, together with the transformation matrix $P \in \mathbb{R}^{3\times 3}$ so that $PA = E_A$.

$$[A\,|\,I] = \begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 1 & 1 & | & 0 & 1 & 0 \\ 1 & 2 & | & 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_3 \leftarrow R_3 - R_1]{R_2 \leftarrow R_2 - R_1} \begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & | & -1 & 1 & 0 \\ 0 & 2 & | & -1 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - 2R_2} \begin{bmatrix} 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & | & -1 & 1 & 0 \\ 0 & 0 & | & 1 & -2 & 1 \end{bmatrix} = [E_A\,|\,P]$$

Therefore, we have two basic columns in A, which forms the basis of $\mathcal{R}(A)$, i.e.,

$$\boxed{\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}}.$$

Clearly, $\boxed{\operatorname{rank}(A) = 2}$.

(b) (5 pts) What is $\mathcal{N}(A)$ in this case?

Answer: The dimension of the input space is 2 since it is $\mathbb{R}^2$. Now from the rank-nullity theorem, we have:

$$\operatorname{rank}(A) + \dim \mathcal{N}(A) = 2.$$

But $\operatorname{rank}(A) = 2$. Hence, $\dim \mathcal{N}(A) = 0$. This means that $\boxed{\mathcal{N}(A) = \{\mathbf{0}\} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}}$.

(c) (5 pts) Find a basis of $\mathcal{R}(A^T)$.

Answer: The nonzero rows of E_A form the basis of $\mathcal{R}(A^T)$. Thus, the answer is:

$$\boxed{\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}}.$$

(d) (5 pts) Find a basis of $\mathcal{N}(A^T)$.

Answer: The row vectors of P corresponding to the zero rows of E_A form the basis of $\mathcal{N}(A^T)$. Thus, the answer is:

$$\boxed{\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}}.$$

Problem 3 (20 pts) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transform defined by $T(x, y) = \big(2x + y, x + 2y\big)$.

Consider the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and the basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

Let $U = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ be the matrix representing this basis.

(a) (5 pts) Determine $[T]_{\mathcal{B}}$ and $[\mathbf{v}]_{\mathcal{B}}$.

Answer: Let A be the matrix associated with this linear transformation T. Then,

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ x + 2y \end{bmatrix} \Longrightarrow A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Now, we have

$$[T]_{\mathcal{B}} = U^{-1} A U = \frac{1}{2} \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \boxed{\frac{1}{2} \begin{bmatrix} 3 & 3 \\ 1 & 5 \end{bmatrix}}.$$

As for $[\mathbf{v}]_{\mathcal{B}}$, it is easy to get:

$$[\mathbf{v}]_{\mathcal{B}} = U^{-1} \mathbf{v} = \frac{1}{2} \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \boxed{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}.$$

(b) (5 pts) Compute $[T(\mathbf{v})]_{\mathcal{B}}$ and verify that $[T]_{\mathcal{B}} [\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$.

Answer: In this case, $T(\mathbf{v}) = A\mathbf{v}$. So,

$$[T(\mathbf{v})]_{\mathcal{B}} = [A\mathbf{v}]_{\mathcal{B}} = U^{-1} A \mathbf{v} = \frac{1}{2} \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \boxed{\frac{1}{2} \begin{bmatrix} 3 \\ 5 \end{bmatrix}}.$$

On the other hand, using the results of Part (a), we have:

$$[T]_{\mathcal{B}} [\mathbf{v}]_{\mathcal{B}} = \frac{1}{2} \begin{bmatrix} 3 & 3 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \boxed{\frac{1}{2} \begin{bmatrix} 3 \\ 5 \end{bmatrix}}.$$

So, they surely agree.

(c) (5 pts) Now, let a new basis in $\mathbb{R}^2$ be

$$\widetilde{\mathcal{B}} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

Let $V = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ be the matrix representing this basis. Now determine the change of basis matrix $[I]_{\mathcal{B}\widetilde{\mathcal{B}}}$.

Answer: The matrix we want to compute is:

$$[I]_{\mathcal{B}\widetilde{\mathcal{B}}} = V^{-1}U = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \boxed{\frac{1}{2}\begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix}}.$$

(d) (5 pts) determine $[T]_{\mathcal{B}\widetilde{\mathcal{B}}}$ and demonstrate that $[I]_{\mathcal{B}\widetilde{\mathcal{B}}}[T]_{\widetilde{\mathcal{B}}} = [T]_{\mathcal{B}\widetilde{\mathcal{B}}}$.

Answer: This can be easily done as follows:

$$[T]_{\mathcal{B}\widetilde{\mathcal{B}}} = V^{-1}AU = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \boxed{\frac{1}{2}\begin{bmatrix} 3 & 9 \\ -1 & 1 \end{bmatrix}}.$$

$$[I]_{\mathcal{B}\widetilde{\mathcal{B}}}[T]_{\mathcal{B}}\frac{1}{2}\begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix} \cdot \frac{1}{2}\begin{bmatrix} 3 & 3 \\ 1 & 5 \end{bmatrix} = \boxed{\frac{1}{2}\begin{bmatrix} 3 & 9 \\ -1 & 1 \end{bmatrix}}.$$

So, they surely agree.

Problem 4 (20 pts) Consider the following matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

(a) (5 pts) Compute the Frobenius norm $\|A\|_F$.

Answer: Using the definition of the Frobenius norm, we have

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{1^2 + 1^2 + 1^2 + (-1)^2 + 0^2 + 1^2} = \boxed{\sqrt{5}}.$$

(b) (4 pts) Compute the 1-norm $\|A\|_1$.

Answer: Using the definition of the 1-norm, we have

$$\|A\|_1 = \max_j \|\boldsymbol{A}_{*j}\|_1 = \max\{3, 2\} = \boxed{3}.$$

(c) (7 pts) Compute the 2-norm $\|A\|_2$.

Answer: We need to compute the eigenvalue of $A^T A$ because $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$. The characteristic equation of $A^T A$ is:

$$\det(A^T A - \lambda I) = \det \begin{bmatrix} 3-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix} = (3-\lambda)(2-\lambda) = 0.$$

Therefore, clearly, $\lambda_{\max} = 3$. Thus, we have $\boxed{\|A\|_2 = \sqrt{3}}$.

(d) (4 pts) Compute the ∞-norm $\|A\|_\infty$.

Answer: Using the definition of the ∞-norm, we have

$$\|A\|_\infty = \max_i \|\boldsymbol{A}_{i*}\|_1 = \max\{2, 1, 2\} = \boxed{2}.$$

Problem 5 (20 pts) Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an *orthonormal basis* of an inner product space $\mathcal{V}$ with $\dim \mathcal{V} = n$.

(a) (5pts) Let $\boldsymbol{x} \in \mathcal{V}$ be any vector in $\mathcal{V}$. Express $\boldsymbol{x}$ as a linear combination of the basis set $\mathcal{B}$, i.e., determine the linear combination coefficients, $\alpha_1, \ldots, \alpha_n$ such that $\boldsymbol{x} = \alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n$.

Answer: Taking an inner product of the both sides of the above equation with $\mathbf{u}_j$ $(1 \le j \le n)$ gives us:

$$\begin{aligned} \langle \mathbf{u}_j, \boldsymbol{x} \rangle &= \langle \mathbf{u}_j, \alpha_1 \mathbf{u}_1 + \cdots + \alpha_j \mathbf{u}_j + \cdots + \alpha_n \mathbf{u}_n \rangle \\ &= \langle \mathbf{u}_j, \alpha_1 \mathbf{u}_1 \rangle + \cdots + \langle \mathbf{u}_j, \alpha_j \mathbf{u}_j \rangle + \cdots + \langle \mathbf{u}_j, \alpha_n \mathbf{u}_n \rangle \\ &= \alpha_1 \langle \mathbf{u}_j, \mathbf{u}_1 \rangle + \cdots + \alpha_j \langle \mathbf{u}_j, \mathbf{u}_j \rangle + \cdots + \alpha_n \langle \mathbf{u}_j, \mathbf{u}_n \rangle \\ &= \alpha_j, \end{aligned}$$

since $\langle \mathbf{u}_j, \mathbf{u}_k \rangle = \delta_{jk}$. This holds for all $j = 1, \ldots, n$. So, we have

$$\boxed{\boldsymbol{x} = \langle \mathbf{u}_1, \boldsymbol{x} \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{u}_n, \boldsymbol{x} \rangle \mathbf{u}_n}.$$

(b) (5 pts) Prove the Pythagorean theorem:

$$\|\boldsymbol{x}\|^2 = |\alpha_1|^2 + \cdots + |\alpha_n|^2.$$

Answer: Using the definition of the norm induced from the inner product, we have:

$$\begin{aligned} \|\boldsymbol{x}\|^2 &= \langle \boldsymbol{x}, \boldsymbol{x} \rangle \\ &= \langle \alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n, \alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n \rangle \\ &= \langle \alpha_1 \mathbf{u}_1, \alpha_1 \mathbf{u}_1 \rangle + \cdots + \langle \alpha_n \mathbf{u}_n, \alpha_n \mathbf{u}_n \rangle \quad \text{because of the orthogonality} \\ &= \|\alpha_1 \mathbf{u}_1\|^2 + \cdots + \|\alpha_n \mathbf{u}_n\|^2 \\ &= |\alpha_1|^2 \|\mathbf{u}_1\|^2 + \cdots + |\alpha_n|^2 \|\mathbf{u}_n\|^2 \\ &= |\alpha_1|^2 + \cdots + |\alpha_n|^2 \quad \text{because they have a unit length.} \end{aligned}$$

(c) (5 pts) Suppose $n = 3$ and $\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$, which are not orthonormal. Make this set orthonormal.

Answer: Let $\mathcal{B} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ be that initial basis. Now, using the Gram-Schmidt procedure, we make the basis vectors orthonormal as follows.

$$\mathbf{u}_1 = \frac{\boldsymbol{x}_1}{\|\boldsymbol{x}_1\|} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

because $\boldsymbol{x}_1$ is already of unit norm.

$$\widetilde{\mathbf{u}}_2 = \boldsymbol{x}_2 - \langle \mathbf{u}_1, \boldsymbol{x}_2 \rangle \mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

which is also of unit norm. So, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Now,

$$\widetilde{\mathbf{u}}_3 = \boldsymbol{x}_3 - \langle \mathbf{u}_1, \boldsymbol{x}_3 \rangle \mathbf{u}_1 - \langle \mathbf{u}_2, \boldsymbol{x}_3 \rangle \mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

which is again of unit norm. Hence, we have the orthonormal set derived from $\mathcal{B}$ as follows:

$$\boxed{\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}}.$$

(d) (5 pts) Expand a vector $\boldsymbol{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ with respect to the orthonormal vectors derived in Part (c).

Answer: This can be easily done because $\boldsymbol{x} = \langle \mathbf{u}_1, \boldsymbol{x} \rangle \mathbf{u}_1 \langle \mathbf{u}_2, \boldsymbol{x} \rangle \mathbf{u}_2 + \langle \mathbf{u}_3, \boldsymbol{x} \rangle \mathbf{u}_3$. The only thing to do is to compute these three inner products, which are in fact, 1, −2, and 1, respectively. Thus, we have

$$\boxed{\boldsymbol{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{u}_1 - 2 \cdot \mathbf{u}_2 + 1 \cdot \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}.$$