MATH 311

Topics in Applied Mathematics

Lecture 20:
The Gram-Schmidt process (continued).
Review for Test 2.

Orthogonal sets

Let V be a vector space with an inner product.

Definition. Nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in V$ form an **orthogonal set** if they are orthogonal to each other: $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for $i \neq j$.

If, in addition, all vectors are of unit norm, $\|\mathbf{v}_i\| = 1$, then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is called an **orthonormal set**.

Theorem Any orthogonal set is linearly independent.

Orthogonal projection

Theorem Let V be a finite-dimensional inner product space and V_0 be a subspace of V. Then any vector $\mathbf{x} \in V$ is uniquely represented as $\mathbf{x} = \mathbf{p} + \mathbf{o}$, where $\mathbf{p} \in V_0$ and $\mathbf{o} \perp V_0$.

The component $\mathbf{p}$ is the **orthogonal projection** of the vector $\mathbf{x}$ onto the subspace V_0. The distance from $\mathbf{x}$ to the subspace V_0 is $\|\mathbf{o}\|$.

If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for V_0 then

$$\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 + \frac{\langle \mathbf{x}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 + \dots + \frac{\langle \mathbf{x}, \mathbf{v}_n \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle}\mathbf{v}_n.$$

The Gram-Schmidt orthogonalization process

Let V be a vector space with an inner product. Suppose $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ is a basis for V. Let

$$\mathbf{v}_1 = \mathbf{x}_1,$$

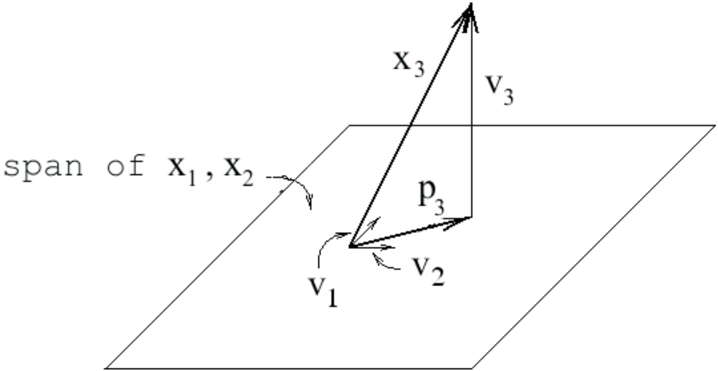
$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1,$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\mathbf{v}_n = \mathbf{x}_n - \frac{\langle \mathbf{x}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \cdots - \frac{\langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle} \mathbf{v}_{n-1}.$$

Then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is an orthogonal basis for V.



span of x_1, x_2
x_3
v_3
p_3
v_1
v_2

Orthogonalization / Normalization

An alternative form of the Gram-Schmidt process combines orthogonalization with normalization.

Suppose $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ is a basis for an inner product space V. Let

$$\mathbf{v}_1 = \mathbf{x}_1, \quad \mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|},$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \langle \mathbf{x}_2, \mathbf{w}_1 \rangle \mathbf{w}_1, \quad \mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|},$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \langle \mathbf{x}_3, \mathbf{w}_1 \rangle \mathbf{w}_1 - \langle \mathbf{x}_3, \mathbf{w}_2 \rangle \mathbf{w}_2, \quad \mathbf{w}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|},$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\mathbf{v}_n = \mathbf{x}_n - \langle \mathbf{x}_n, \mathbf{w}_1 \rangle \mathbf{w}_1 - \cdots - \langle \mathbf{x}_n, \mathbf{w}_{n-1} \rangle \mathbf{w}_{n-1},$$
$$\mathbf{w}_n = \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}.$$

Then $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ is an orthonormal basis for V.

Problem. Find the distance from the point $\mathbf{y} = (0,0,0,1)$ to the subspace $\Pi \subset \mathbb{R}^4$ spanned by vectors $\mathbf{x}_1 = (1,-1,1,-1)$, $\mathbf{x}_2 = (1,1,3,-1)$, and $\mathbf{x}_3 = (-3,7,1,3)$.

Let us apply the Gram-Schmidt process to vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{y}$. We should obtain an orthogonal system $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$. The desired distance will be $|\mathbf{v}_4|$.

$\mathbf{x}_1 = (1, -1, 1, -1),\ \ \mathbf{x}_2 = (1, 1, 3, -1),$
$\mathbf{x}_3 = (-3, 7, 1, 3),\ \ \mathbf{y} = (0, 0, 0, 1).$

$$\mathbf{v}_1 = \mathbf{x}_1 = (1, -1, 1, -1),$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 = (1, 1, 3, -1) - \frac{4}{4}(1, -1, 1, -1)$$

$$= (0, 2, 2, 0),$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2$$

$$= (-3, 7, 1, 3) - \frac{-12}{4}(1, -1, 1, -1) - \frac{16}{8}(0, 2, 2, 0)$$

$$= (0, 0, 0, 0).$$

The Gram-Schmidt process can be used to check linear independence of vectors!

The vector $\mathbf{x}_3$ is a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$.
Π is a plane, not a 3-dimensional subspace.
We should orthogonalize vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}$.

$$\mathbf{v}_4 = \mathbf{y} - \frac{\langle \mathbf{y}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{y}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2$$

$$= (0,0,0,1) - \frac{-1}{4}(1,-1,1,-1) - \frac{0}{8}(0,2,2,0)$$

$$= (1/4, -1/4, 1/4, 3/4).$$

$$|\mathbf{v}_4| = \left| \left(\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{3}{4} \right) \right| = \frac{1}{4} \left| (1,-1,1,3) \right| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2}.$$

Problem. Find the distance from the point $\mathbf{z} = (0, 0, 1, 0)$ to the plane Π that passes through the point $\mathbf{x}_0 = (1, 0, 0, 0)$ and is parallel to the vectors $\mathbf{v}_1 = (1, -1, 1, -1)$ and $\mathbf{v}_2 = (0, 2, 2, 0)$.

The plane Π is not a subspace of $\mathbb{R}^4$ as it does not pass through the origin. Let $\Pi_0 = \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2)$. Then $\Pi = \Pi_0 + \mathbf{x}_0$.

Hence the distance from the point $\mathbf{z}$ to the plane Π is the same as the distance from the point $\mathbf{z} - \mathbf{x}_0$ to the plane Π_0.

We shall apply the Gram-Schmidt process to vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{z} - \mathbf{x}_0$. This will yield an orthogonal system $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. The desired distance will be $|\mathbf{w}_3|$.

$$\mathbf{v}_1 = (1,-1,1,-1),\ \mathbf{v}_2 = (0,2,2,0),\ \mathbf{z}-\mathbf{x}_0 = (-1,0,1,0).$$

$$\mathbf{w}_1 = \mathbf{v}_1 = (1,-1,1,-1),$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1\rangle}{\langle \mathbf{w}_1, \mathbf{w}_1\rangle}\mathbf{w}_1 = \mathbf{v}_2 = (0,2,2,0) \quad \text{as } \mathbf{v}_2 \perp \mathbf{v}_1.$$

$$\mathbf{w}_3 = (\mathbf{z}-\mathbf{x}_0) - \frac{\langle \mathbf{z}-\mathbf{x}_0, \mathbf{w}_1\rangle}{\langle \mathbf{w}_1, \mathbf{w}_1\rangle}\mathbf{w}_1 - \frac{\langle \mathbf{z}-\mathbf{x}_0, \mathbf{w}_2\rangle}{\langle \mathbf{w}_2, \mathbf{w}_2\rangle}\mathbf{w}_2$$

$$= (-1,0,1,0) - \frac{0}{4}(1,-1,1,-1) - \frac{2}{8}(0,2,2,0)$$

$$= (-1,-1/2,1/2,0).$$

$$|\mathbf{w}_3| = \left|\left(-1,-\frac{1}{2},\frac{1}{2},0\right)\right| = \frac{1}{2}\,|(-2,-1,1,0)| = \frac{\sqrt{6}}{2} = \sqrt{\frac{3}{2}}.$$

Topics for Test 2

Coordinates and linear transformations (Leon 3.5, 4.1–4.3)

- Coordinates relative to a basis
- Change of basis, transition matrix
- Matrix transformations
- Matrix of a linear mapping

Eigenvalues and eigenvectors (Leon 6.1, 6.3)

- Eigenvalues, eigenvectors, eigenspaces
- Characteristic polynomial
- Diagonalization

Orthogonality (Leon 5.1–5.6)

- Inner products and norms
- Orthogonal complement
- Least squares problems
- The Gram-Schmidt orthogonalization process

Sample problems for Test 2

Problem 1 (15 pts.) Let $\mathcal{M}_{2,2}(\mathbb{R})$ denote the vector space of 2×2 matrices with real entries. Consider a linear operator $L : \mathcal{M}_{2,2}(\mathbb{R}) \to \mathcal{M}_{2,2}(\mathbb{R})$ given by

$$L\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Find the matrix of the operator L with respect to the basis

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Problem 2 (30 pts.) Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

(i) Find all eigenvalues of the matrix A.
(ii) For each eigenvalue of A, find an associated eigenvector.
(iii) Is the matrix A diagonalizable? Explain.
(iv) Find all eigenvalues of the matrix A^2.

Problem 3 (20 pts.) Find a linear polynomial which is the best least squares fit to the following data:

x	-2	-1	0	1	2
$f(x)$	-3	-2	1	2	5

Problem 4 (25 pts.) Let V be a subspace of $\mathbb{R}^4$ spanned by the vectors $\mathbf{x}_1 = (1, 1, 1, 1)$ and $\mathbf{x}_2 = (1, 0, 3, 0)$.

(i) Find an orthonormal basis for V.
(ii) Find an orthonormal basis for the orthogonal complement $V^{\perp}$.

Bonus Problem 5 (15 pts.) Let $L : V \to W$ be a linear mapping of a finite-dimensional vector space V to a vector space W. Show that

$$\dim \mathrm{Range}(L) + \dim \ker(L) = \dim V.$$

Problem 1. Let $\mathcal{M}_{2,2}(\mathbb{R})$ denote the vector space of 2×2 matrices with real entries. Consider a linear operator $L : \mathcal{M}_{2,2}(\mathbb{R}) \to \mathcal{M}_{2,2}(\mathbb{R})$ given by

$$L\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Find the matrix of the operator L with respect to the basis

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Let M_L denote the desired matrix.

By definition, M_L is a 4×4 matrix whose columns are coordinates of the matrices $L(E_1), L(E_2), L(E_3), L(E_4)$ with respect to the basis E_1, E_2, E_3, E_4.

$$L(E_1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = 1E_1 + 2E_2 + 0E_3 + 0E_4,$$

$$L(E_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 0 & 0 \end{pmatrix} = 3E_1 + 4E_2 + 0E_3 + 0E_4,$$

$$L(E_3) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} = 0E_1 + 0E_2 + 1E_3 + 2E_4,$$

$$L(E_4) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix} = 0E_1 + 0E_2 + 3E_3 + 4E_4.$$

It follows that

$$M_L = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 2 & 4 \end{pmatrix}.$$

Thus the relation

$$\begin{pmatrix} x_1 & y_1 \\ z_1 & w_1 \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

is equivalent to the relation

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}.$$

Problem 2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

(i) Find all eigenvalues of the matrix A.

The eigenvalues of A are roots of the characteristic equation $\det(A - \lambda I) = 0$. We obtain that

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 1 & 1-\lambda & 1 \\ 0 & 2 & 1-\lambda \end{vmatrix}$$

$$= (1-\lambda)^3 - 2(1-\lambda) - 2(1-\lambda) = (1-\lambda)\big((1-\lambda)^2 - 4\big)$$

$$= (1-\lambda)\big((1-\lambda) - 2\big)\big((1-\lambda) + 2\big) = -(\lambda - 1)(\lambda + 1)(\lambda - 3).$$

Hence the matrix A has three eigenvalues: -1, 1, and 3.

Problem 2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

(ii) For each eigenvalue of A, find an associated eigenvector.

An eigenvector $\mathbf{v} = (x, y, z)$ of the matrix A associated with an eigenvalue λ is a nonzero solution of the vector equation

$$(A - \lambda I)\mathbf{v} = \mathbf{0} \iff \begin{pmatrix} 1-\lambda & 2 & 0 \\ 1 & 1-\lambda & 1 \\ 0 & 2 & 1-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

To solve the equation, we convert the matrix $A - \lambda I$ to reduced row echelon form.

First consider the case $\lambda = -1$. The row reduction yields

$$A + I = \begin{pmatrix} 2 & 2 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} \color{red}{1} & 0 & -1 \\ 0 & \color{red}{1} & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hence

$$(A + I)\mathbf{v} = \mathbf{0} \iff \begin{cases} x - z = 0, \\ y + z = 0. \end{cases}$$

The general solution is $x = t$, $y = -t$, $z = t$, where $t \in \mathbb{R}$. In particular, $\mathbf{v}_1 = (1, -1, 1)$ is an eigenvector of A associated with the eigenvalue -1.

Secondly, consider the case $\lambda = 1$. The row reduction yields

$$A - I = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 2 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hence

$$(A - I)\mathbf{v} = \mathbf{0} \quad \Longleftrightarrow \quad \begin{cases} x + z = 0, \\ y = 0. \end{cases}$$

The general solution is $x = -t$, $y = 0$, $z = t$, where $t \in \mathbb{R}$. In particular, $\mathbf{v}_2 = (-1, 0, 1)$ is an eigenvector of A associated with the eigenvalue 1.

Finally, consider the case $\lambda = 3$. The row reduction yields

$$A - 3I = \begin{pmatrix} -2 & 2 & 0 \\ 1 & -2 & 1 \\ 0 & 2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 \\ 1 & -2 & 1 \\ 0 & 2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 2 & -2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hence

$$(A - 3I)\mathbf{v} = \mathbf{0} \iff \begin{cases} x - z = 0, \\ y - z = 0. \end{cases}$$

The general solution is $x = t$, $y = t$, $z = t$, where $t \in \mathbb{R}$. In particular, $\mathbf{v}_3 = (1, 1, 1)$ is an eigenvector of A associated with the eigenvalue 3.

Problem 2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

(iii) Is the matrix A diagonalizable? Explain.

The matrix A is diagonalizable, i.e., there exists a basis for $\mathbb{R}^3$ formed by its eigenvectors.

Namely, the vectors $\mathbf{v}_1 = (1, -1, 1)$, $\mathbf{v}_2 = (-1, 0, 1)$, and $\mathbf{v}_3 = (1, 1, 1)$ are eigenvectors of the matrix A belonging to distinct eigenvalues. Therefore these vectors are linearly independent. It follows that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is a basis for $\mathbb{R}^3$.

Alternatively, the existence of a basis for $\mathbb{R}^3$ consisting of eigenvectors of A already follows from the fact that the matrix A has three distinct eigenvalues.

Problem 2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

(iv) Find all eigenvalues of the matrix A^2.

We have that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

The columns of U are eigenvectors of A belonging to the eigenvalues -1, 1, and 3, respectively.

Then $A^2 = UBU^{-1}UBU^{-1} = UB^2U^{-1}$, that is, the matrix A^2 is similar to the diagonal matrix $B^2 = \mathrm{diag}(1, 1, 9)$.

Similar matrices have the same characteristic polynomial, hence they have the same eigenvalues. Thus the eigenvalues of A^2 are the same as the eigenvalues of B^2: 1 and 9.

Problem 3. Find a linear polynomial which is the best least squares fit to the following data:

x	-2	-1	0	1	2
$f(x)$	-3	-2	1	2	5

We are looking for a function $f(x) = c_1 + c_2 x$, where c_1, c_2 are unknown coefficients. The data of the problem give rise to an overdetermined system of linear equations in variables c_1 and c_2:

$$\begin{cases} c_1 - 2c_2 = -3, \\ c_1 - c_2 = -2, \\ c_1 = 1, \\ c_1 + c_2 = 2, \\ c_1 + 2c_2 = 5. \end{cases}$$

This system is inconsistent.

We can represent the system as a matrix equation $A\mathbf{c} = \mathbf{y}$, where

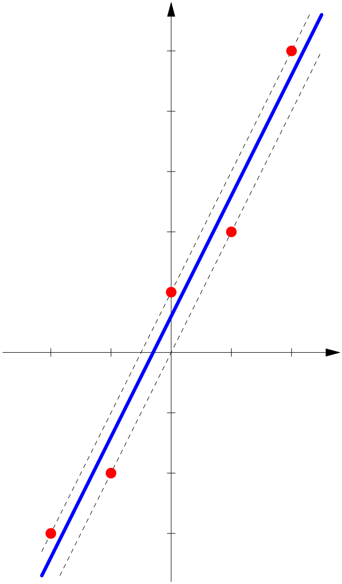
$$A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} -3 \\ -2 \\ 1 \\ 2 \\ 5 \end{pmatrix}.$$

The least squares solution $\mathbf{c}$ of the above system is a solution of the normal system $A^T A\mathbf{c} = A^T\mathbf{y}$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ -2 \\ 1 \\ 2 \\ 5 \end{pmatrix}$$

$$\iff \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 20 \end{pmatrix} \iff \begin{cases} c_1 = 3/5 \\ c_2 = 2 \end{cases}$$

Thus the function $f(x) = \frac{3}{5} + 2x$ is the best least squares fit to the above data among linear polynomials.



Problem 4. Let V be a subspace of $\mathbb{R}^4$ spanned by the vectors $\mathbf{x}_1 = (1,1,1,1)$ and $\mathbf{x}_2 = (1,0,3,0)$.

(i) Find an orthonormal basis for V.

First we apply the Gram-Schmidt orthogonalization process to vectors $\mathbf{x}_1, \mathbf{x}_2$ and obtain an orthogonal basis $\mathbf{v}_1, \mathbf{v}_2$ for the subspace V:

$$\mathbf{v}_1 = \mathbf{x}_1 = (1,1,1,1),$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = (1,0,3,0) - \frac{4}{4}(1,1,1,1) = (0,-1,2,-1).$$

Then we normalize vectors $\mathbf{v}_1, \mathbf{v}_2$ to obtain an orthonormal basis $\mathbf{w}_1, \mathbf{w}_2$ for V:

$$\|\mathbf{v}_1\| = 2 \implies \mathbf{w}_1 = \tfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \tfrac{1}{2}(1,1,1,1)$$

$$\|\mathbf{v}_2\| = \sqrt{6} \implies \mathbf{w}_2 = \tfrac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \tfrac{1}{\sqrt{6}}(0,-1,2,-1)$$

Problem 4. Let V be a subspace of $\mathbb{R}^4$ spanned by the vectors $\mathbf{x}_1 = (1,1,1,1)$ and $\mathbf{x}_2 = (1,0,3,0)$.

(ii) Find an orthonormal basis for the orthogonal complement $V^\perp$.

Since the subspace V is spanned by vectors $(1,1,1,1)$ and $(1,0,3,0)$, it is the row space of the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 3 & 0 \end{pmatrix}.$$

Then the orthogonal complement $V^\perp$ is the nullspace of A. To find the nullspace, we convert the matrix A to reduced row echelon form:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 3 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix}.$$

Hence a vector $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ belongs to $V^\perp$ if and only if

$$\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff \quad \begin{cases} x_1 + 3x_3 = 0 \\ x_2 - 2x_3 + x_4 = 0 \end{cases} \quad \iff \quad \begin{cases} x_1 = -3x_3 \\ x_2 = 2x_3 - x_4 \end{cases}$$

The general solution of the system is $(x_1, x_2, x_3, x_4) =$
$= (-3t, 2t - s, t, s) = t(-3, 2, 1, 0) + s(0, -1, 0, 1)$, where $t, s \in \mathbb{R}$.

It follows that $V^\perp$ is spanned by vectors $\mathbf{x}_3 = (0, -1, 0, 1)$ and $\mathbf{x}_4 = (-3, 2, 1, 0)$.

The vectors $\mathbf{x}_3 = (0, -1, 0, 1)$ and $\mathbf{x}_4 = (-3, 2, 1, 0)$ form a basis for the subspace $V^\perp$.

It remains to orthogonalize and normalize this basis:

$$\mathbf{v}_3 = \mathbf{x}_3 = (0, -1, 0, 1),$$

$$\mathbf{v}_4 = \mathbf{x}_4 - \frac{\mathbf{x}_4 \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3}\,\mathbf{v}_3 = (-3, 2, 1, 0) - \frac{-2}{2}(0, -1, 0, 1)$$
$$= (-3, 1, 1, 1),$$

$$\|\mathbf{v}_3\| = \sqrt{2} \implies \mathbf{w}_3 = \tfrac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \tfrac{1}{\sqrt{2}}(0, -1, 0, 1),$$

$$\|\mathbf{v}_4\| = \sqrt{12} = 2\sqrt{3} \implies \mathbf{w}_4 = \tfrac{\mathbf{v}_4}{\|\mathbf{v}_4\|} = \tfrac{1}{2\sqrt{3}}(-3, 1, 1, 1).$$

Thus the vectors $\mathbf{w}_3 = \frac{1}{\sqrt{2}}(0, -1, 0, 1)$ and $\mathbf{w}_4 = \frac{1}{2\sqrt{3}}(-3, 1, 1, 1)$ form an orthonormal basis for $V^\perp$.

Problem 4. Let V be a subspace of $\mathbb{R}^4$ spanned by the vectors $\mathbf{x}_1 = (1, 1, 1, 1)$ and $\mathbf{x}_2 = (1, 0, 3, 0)$.
(i) Find an orthonormal basis for V.
(ii) Find an orthonormal basis for the orthogonal complement $V^\perp$.

Alternative solution: First we extend the set $\mathbf{x}_1, \mathbf{x}_2$ to a basis $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ for $\mathbb{R}^4$. Then we orthogonalize and normalize the latter. This yields an orthonormal basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ for $\mathbb{R}^4$.

By construction, $\mathbf{w}_1, \mathbf{w}_2$ is an orthonormal basis for V.
It follows that $\mathbf{w}_3, \mathbf{w}_4$ is an orthonormal basis for $V^\perp$.

The set $\mathbf{x}_1 = (1, 1, 1, 1)$, $\mathbf{x}_2 = (1, 0, 3, 0)$ can be extended to a basis for $\mathbb{R}^4$ by adding two vectors from the standard basis.

For example, we can add vectors $\mathbf{e}_3 = (0, 0, 1, 0)$ and $\mathbf{e}_4 = (0, 0, 0, 1)$. To show that $\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_3, \mathbf{e}_4$ is indeed a basis for $\mathbb{R}^4$, we check that the matrix whose rows are these vectors is nonsingular:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1 \neq 0.$$

To orthogonalize the basis $\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_3, \mathbf{e}_4$, we apply the Gram-Schmidt process:

$$\mathbf{v}_1 = \mathbf{x}_1 = (1, 1, 1, 1),$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = (1, 0, 3, 0) - \tfrac{4}{4}(1, 1, 1, 1) = (0, -1, 2, -1),$$

$$\mathbf{v}_3 = \mathbf{e}_3 - \frac{\mathbf{e}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{e}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = (0, 0, 1, 0) - \tfrac{1}{4}(1, 1, 1, 1) -$$

$$- \tfrac{2}{6}(0, -1, 2, -1) = \left(-\tfrac{1}{4}, \tfrac{1}{12}, \tfrac{1}{12}, \tfrac{1}{12}\right) = \tfrac{1}{12}(-3, 1, 1, 1),$$

$$\mathbf{v}_4 = \mathbf{e}_4 - \frac{\mathbf{e}_4 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{e}_4 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \frac{\mathbf{e}_4 \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3}\mathbf{v}_3 = (0, 0, 0, 1) -$$

$$- \tfrac{1}{4}(1, 1, 1, 1) - \tfrac{-1}{6}(0, -1, 2, -1) - \tfrac{1/12}{1/12} \cdot \tfrac{1}{12}(-3, 1, 1, 1) =$$

$$= \left(0, -\tfrac{1}{2}, 0, \tfrac{1}{2}\right) = \tfrac{1}{2}(0, -1, 0, 1).$$

It remains to normalize vectors $\mathbf{v}_1 = (1, 1, 1, 1)$, $\mathbf{v}_2 = (0, -1, 2, -1)$, $\mathbf{v}_3 = \frac{1}{12}(-3, 1, 1, 1)$, $\mathbf{v}_4 = \frac{1}{2}(0, -1, 0, 1)$:

$$\|\mathbf{v}_1\| = 2 \implies \mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{2}(1, 1, 1, 1)$$

$$\|\mathbf{v}_2\| = \sqrt{6} \implies \mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{1}{\sqrt{6}}(0, -1, 2, -1)$$

$$\|\mathbf{v}_3\| = \frac{1}{\sqrt{12}} = \frac{1}{2\sqrt{3}} \implies \mathbf{w}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{1}{2\sqrt{3}}(-3, 1, 1, 1)$$

$$\|\mathbf{v}_4\| = \frac{1}{\sqrt{2}} \implies \mathbf{w}_4 = \frac{\mathbf{v}_4}{\|\mathbf{v}_4\|} = \frac{1}{\sqrt{2}}(0, -1, 0, 1)$$

Thus $\mathbf{w}_1, \mathbf{w}_2$ is an orthonormal basis for V while $\mathbf{w}_3, \mathbf{w}_4$ is an orthonormal basis for $V^\perp$.

Bonus Problem 5. Let $L : V \to W$ be a linear mapping of a finite-dimensional vector space V to a vector space W. Show that $\dim \mathrm{Range}(L) + \dim \ker(L) = \dim V$.

The kernel $\ker(L)$ is a subspace of V. It is finite-dimensional since the vector space V is.

Take a basis $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ for the subspace $\ker(L)$, then extend it to a basis $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ for the entire space V.

Claim Vectors $L(\mathbf{u}_1), L(\mathbf{u}_2), \dots, L(\mathbf{u}_m)$ form a basis for the range of L.

Assuming the claim is proved, we obtain

$$\dim \mathrm{Range}(L) = m, \quad \dim \ker(L) = k, \quad \dim V = k + m.$$

Claim Vectors $L(\mathbf{u}_1), L(\mathbf{u}_2), \dots, L(\mathbf{u}_m)$ form a basis for the range of L.

Proof (spanning): Any vector $\mathbf{w} \in \mathrm{Range}(L)$ is represented as $\mathbf{w} = L(\mathbf{v})$, where $\mathbf{v} \in V$. Then

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k + \beta_1 \mathbf{u}_1 + \beta_2 \mathbf{u}_2 + \cdots + \beta_m \mathbf{u}_m$$

for some $\alpha_i, \beta_j \in \mathbb{R}$. It follows that

$$\mathbf{w} = L(\mathbf{v}) = \alpha_1 L(\mathbf{v}_1) + \cdots + \alpha_k L(\mathbf{v}_k) + \beta_1 L(\mathbf{u}_1) + \cdots + \beta_m L(\mathbf{u}_m)$$

$$= \beta_1 L(\mathbf{u}_1) + \cdots + \beta_m L(\mathbf{u}_m).$$

Note that $L(\mathbf{v}_i) = \mathbf{0}$ since $\mathbf{v}_i \in \ker(L)$.
Thus $\mathrm{Range}(L)$ is spanned by the vectors $L(\mathbf{u}_1), \dots, L(\mathbf{u}_m)$.

Claim Vectors $L(\mathbf{u}_1), L(\mathbf{u}_2), \dots, L(\mathbf{u}_m)$ form a basis for the range of L.

Proof (linear independence): Suppose that

$$t_1 L(\mathbf{u}_1) + t_2 L(\mathbf{u}_2) + \cdots + t_m L(\mathbf{u}_m) = \mathbf{0}$$

for some $t_i \in \mathbb{R}$. Let $\mathbf{u} = t_1\mathbf{u}_1 + t_2\mathbf{u}_2 + \cdots + t_m\mathbf{u}_m$. Since

$$L(\mathbf{u}) = t_1 L(\mathbf{u}_1) + t_2 L(\mathbf{u}_2) + \cdots + t_m L(\mathbf{u}_m) = \mathbf{0},$$

the vector $\mathbf{u}$ belongs to the kernel of L. Therefore $\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_k\mathbf{v}_k$ for some $s_j \in \mathbb{R}$. It follows that

$$t_1\mathbf{u}_1 + t_2\mathbf{u}_2 + \cdots + t_m\mathbf{u}_m - s_1\mathbf{v}_1 - s_2\mathbf{v}_2 - \cdots - s_k\mathbf{v}_k = \mathbf{u} - \mathbf{u} = \mathbf{0}.$$

Linear independence of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}_1, \dots, \mathbf{u}_m$ implies that $t_1 = \cdots = t_m = 0$ (as well as $s_1 = \cdots = s_k = 0$).
Thus the vectors $L(\mathbf{u}_1), L(\mathbf{u}_2), \dots, L(\mathbf{u}_m)$ are linearly independent.