

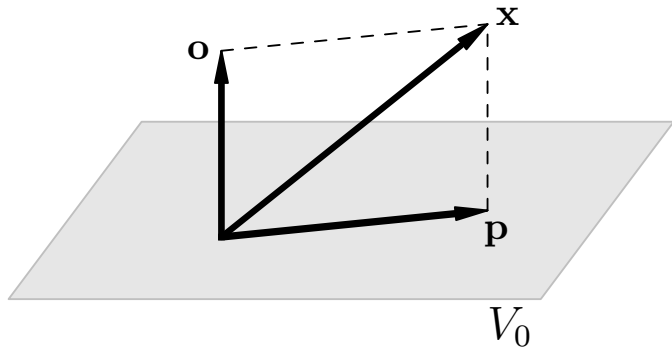
MATH 304

Linear Algebra

Lecture 30:

**The Gram-Schmidt process (continued).
Eigenvalues and eigenvectors.**

Orthogonal projection



Orthogonal projection

Theorem Let V be an inner product space and V_0 be a finite-dimensional 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

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

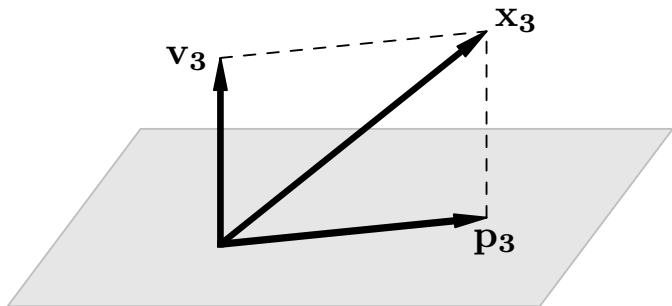
$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1,$$

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

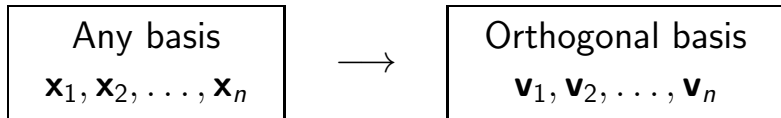
.....

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

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



$$\text{Span}(\mathbf{v}_1, \mathbf{v}_2) = \text{Span}(\mathbf{x}_1, \mathbf{x}_2)$$



Properties of the Gram-Schmidt process:

- $\mathbf{v}_k = \mathbf{x}_k - (\alpha_1 \mathbf{x}_1 + \dots + \alpha_{k-1} \mathbf{x}_{k-1})$, $1 \leq k \leq n$;
- the span of $\mathbf{v}_1, \dots, \mathbf{v}_k$ is the same as the span of $\mathbf{x}_1, \dots, \mathbf{x}_k$;
- \mathbf{v}_k is orthogonal to $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$;
- $\mathbf{v}_k = \mathbf{x}_k - \mathbf{p}_k$, where \mathbf{p}_k is the orthogonal projection of the vector \mathbf{x}_k on the subspace spanned by $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$;
- $\|\mathbf{v}_k\|$ is the distance from \mathbf{x}_k to the subspace spanned by $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$.

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

First we apply the Gram-Schmidt process to vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and obtain an orthogonal basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ for the subspace V . Next we compute the orthogonal projection \mathbf{p} of the vector \mathbf{y} onto V :

$$\mathbf{p} = \frac{\langle \mathbf{y}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 + \frac{\langle \mathbf{y}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 + \frac{\langle \mathbf{y}, \mathbf{v}_3 \rangle}{\langle \mathbf{v}_3, \mathbf{v}_3 \rangle} \mathbf{v}_3.$$

Then the distance from \mathbf{y} to V equals $\|\mathbf{y} - \mathbf{p}\|$.

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

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

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

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

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 \\ = (-3, 7, 1, 3) - \frac{-12}{4}(1, -1, 1, -1) - \frac{16}{8}(0, 2, 2, 0) \\ = (0, 0, 0, 0).$$

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 .

V is a plane, not a 3-dimensional subspace.

We should orthogonalize vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}$.

$$\begin{aligned}\tilde{\mathbf{v}}_3 &= \mathbf{y} - \frac{\langle \mathbf{y}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{y}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 \\ &= (0, 0, 0, 1) - \frac{-1}{4}(1, -1, 1, -1) - \frac{0}{8}(0, 2, 2, 0) \\ &= (1/4, -1/4, 1/4, 3/4).\end{aligned}$$

$$|\tilde{\mathbf{v}}_3| = \left| \left(\frac{1}{4}, -\frac{1}{4}, \frac{1}{4}, \frac{3}{4} \right) \right| = \frac{1}{4} |(1, -1, 1, 3)| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2}.$$

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 $\Pi_0 = \text{Span}(\mathbf{v}_1, \mathbf{v}_2)$. 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 $\Pi - \mathbf{x}_0 = \Pi_0$.

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 $\|\mathbf{w}_3\|$.

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

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

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

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

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

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

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

Problem. Approximate the function $f(x) = e^x$ on the interval $[-1, 1]$ by a quadratic polynomial.

The best approximation would be a polynomial $p(x)$ that minimizes the distance relative to the uniform norm:

$$\|f - p\|_{\infty} = \max_{|x| \leq 1} |f(x) - p(x)|.$$

However there is no analytic way to find such a polynomial. Instead, one can find a “*least squares*” approximation that minimizes the integral norm

$$\|f - p\|_2 = \left(\int_{-1}^1 |f(x) - p(x)|^2 dx \right)^{1/2}.$$

The norm $\| \cdot \|_2$ is induced by the inner product

$$\langle g, h \rangle = \int_{-1}^1 g(x)h(x) dx.$$

Therefore $\|f - p\|_2$ is minimal if p is the orthogonal projection of the function f on the subspace \mathcal{P}_3 of quadratic polynomials.

We should apply the Gram-Schmidt process to the polynomials $1, x, x^2$ which form a basis for \mathcal{P}_3 .

This would yield an orthogonal basis p_0, p_1, p_2 .

Then

$$p(x) = \frac{\langle f, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0(x) + \frac{\langle f, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1(x) + \frac{\langle f, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2(x).$$

Eigenvalues and eigenvectors

Definition. Let A be an $n \times n$ matrix. A number $\lambda \in \mathbb{R}$ is called an **eigenvalue** of the matrix A if $A\mathbf{v} = \lambda\mathbf{v}$ for a nonzero column vector $\mathbf{v} \in \mathbb{R}^n$.

The vector \mathbf{v} is called an **eigenvector** of A belonging to (or associated with) the eigenvalue λ .

Remarks.

- Alternative notation:
eigenvalue = **characteristic value**,
eigenvector = **characteristic vector**.

- The zero vector is never considered an eigenvector.

Example. $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ -2 \end{pmatrix}.$$

Hence $(1, 0)$ is an eigenvector of A belonging to the eigenvalue 2, while $(0, -2)$ is an eigenvector of A belonging to the eigenvalue 3.

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

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Hence $(1, 1)$ is an eigenvector of A belonging to the eigenvalue 1, while $(1, -1)$ is an eigenvector of A belonging to the eigenvalue -1 .

Vectors $\mathbf{v}_1 = (1, 1)$ and $\mathbf{v}_2 = (1, -1)$ form a basis for \mathbb{R}^2 . Consider a linear operator $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by $L(\mathbf{x}) = A\mathbf{x}$. The matrix of L with respect to the basis $\mathbf{v}_1, \mathbf{v}_2$ is $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Let A be an $n \times n$ matrix. Consider a linear operator $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $L(\mathbf{x}) = A\mathbf{x}$.

Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be a nonstandard basis for \mathbb{R}^n and B be the matrix of the operator L with respect to this basis.

Theorem The matrix B is diagonal if and only if vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are eigenvectors of A .

If this is the case, then the diagonal entries of the matrix B are the corresponding eigenvalues of A .

$$A\mathbf{v}_i = \lambda_i\mathbf{v}_i \iff B = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$