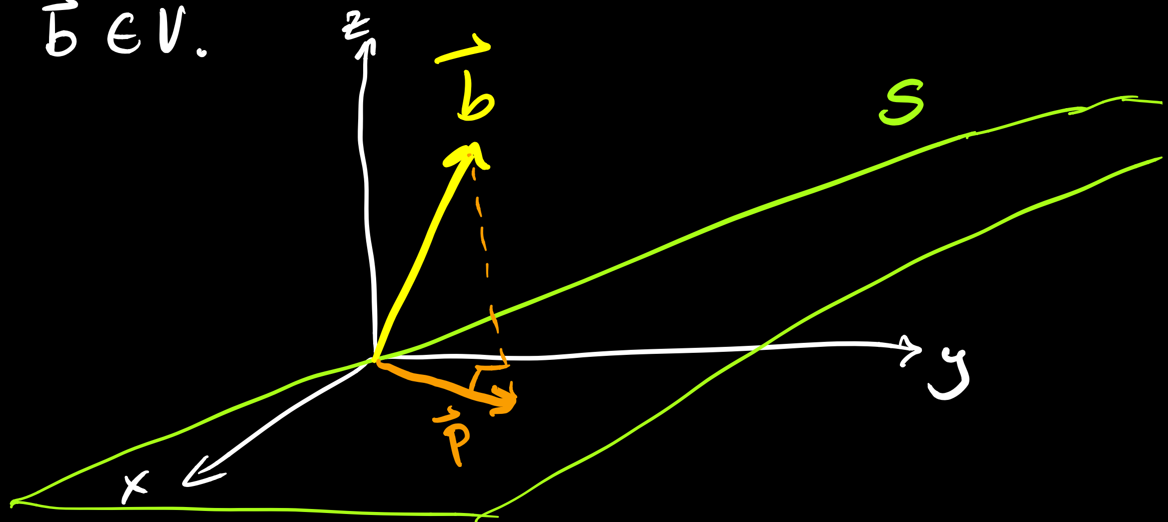
Projections and Least Squares

Given some subspace S of a vector space V, we want to find the best approximation $\vec{p}$ of a vector $\vec{b} \in V$.



To find the projection $\vec{p}$, we want $\vec{x}$ so that

$$A\vec{x} = \vec{b}$$

where $C(A) = S$. Since $C(A) \neq V$, there will be no solution if $\vec{b} \notin C(A)$. The best approximation is $\vec{p}$, which is the vector formed by a linear combination of the columns of A and the coefficients are the solution to the least squares problem.

$$A\vec{x} = \vec{b}$$

$$A^T A \vec{x} = A^T \vec{b}$$

$$\vec{x} = (A^T A)^{-1} A^T \vec{b}$$

least squares solution

Now, $\vec{p} = A\vec{x} = A(A^T A)^{-1} A^T \vec{b}$

and $P = A(A^T A)^{-1} A^T$ is the projection matrix.

<u>Ex</u> Find the projection of $\vec{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$ onto

$$\text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} \right).$$

First we need to make sure the given vectors are linearly independent.

$$\begin{array}{c} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & -1 \end{bmatrix} \end{array}\begin{array}{l} \\ \\ r_3 - r_1 \\ r_4 - 2r_1 \end{array} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & -3 & -3 \end{bmatrix}\begin{array}{l} r_1 - 2r_2 \\ \\ r_3 + r_2 \\ r_4 + 3r_2 \end{array}$$

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

We see that the 1st two columns form a basis of the given subspace, so let

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

Now, we want to find $\vec{x}$, so that

$$A\vec{x} = \vec{b}$$

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

A^TA is invertible if columns of A are linearly indep.

$$\begin{bmatrix} 6 & 5 \\ 5 & 7 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} 6 & 5 \\ 5 & 7 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$= \frac{1}{42-25}\begin{bmatrix} 7 & -5 \\ -5 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$= \frac{1}{17}\begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

Now, $$\vec{p} = \frac{4}{17}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} + \frac{2}{17}\begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 8/17 \\ 2/17 \\ 6/17 \\ 10/17 \end{bmatrix}$$

$$= A\vec{x}$$

Gram-Schmidt

Given a basis of some subspace S, we can use Gram-Schmidt to form an orthonormal basis of S. $\{\vec{q}_1, \ldots, \vec{q}_n\}$ is an orthonormal basis of S if

(a) $\text{span}(\vec{q}_1, \ldots, \vec{q}_n) = S$

(b) $\vec{q}_i^{\,T} \vec{q}_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

Given a set of linearly independent vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ we can make them orthonormal using Gram-Schmidt.

1. Orthogonalize

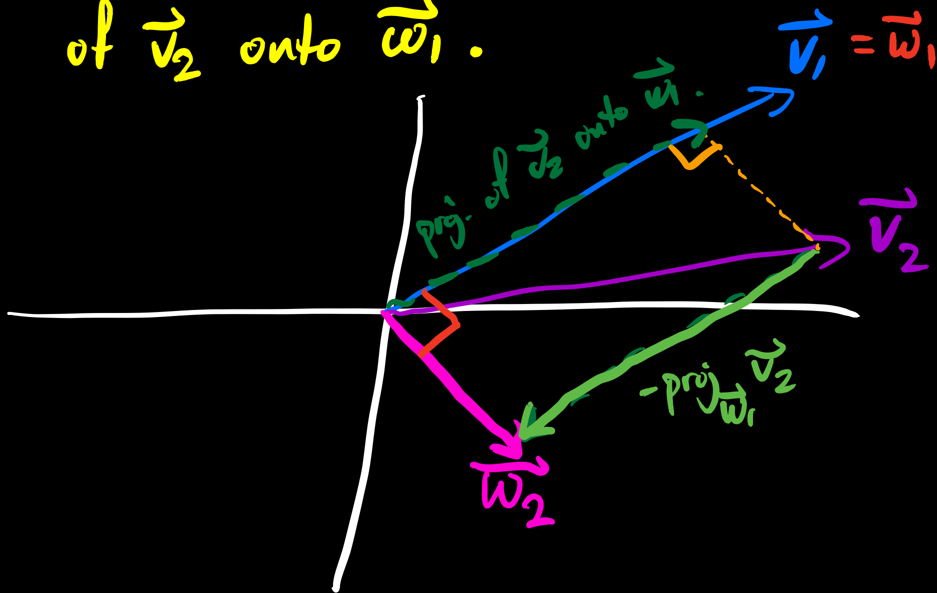
$$\vec{w}_1 = \vec{v}_1$$

$$\vec{w}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2} \vec{w}_1$$

$$= \vec{v}_2 - \vec{w}_1 (\vec{w}_1^{\,T} \vec{w}_1)^{-1} \vec{w}_1^{\,T} \vec{v}_2$$

$A(A^TA)^{-1}A^T\vec{v}_2$ (proj. of $\vec{v}_2$ onto $\vec{w}_1$)

From $\vec{v}_2$, we are removing the projection of $\vec{v}_2$ onto $\vec{w}_1$.



So in each step of the orthogonalization, we remove the projection of $\vec{v}_i$ onto the $\vec{w}_{i-1}, \vec{w}_{i-2}, \cdots, \vec{w}_1$.

$$\vec{w}_i = \vec{v}_i - \sum_{j=1}^{i-1} \vec{w}_j (\vec{w}_j^T \vec{w}_j)^{-1} w_j^T v_i^T$$

2. Normalize $\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_n$, $\vec{q}_i = \frac{1}{\|\vec{w}_i\|} \vec{w}_i$

Ex Find an orthonormal basis of

$$S = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right)$$

The given vectors are linearly independent.

Now, we perform Gram-Schmidt.

1. Orthogonalize

$$\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

$$\vec{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}}{\left\| \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\|^2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{5}{6} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 7/6 \\ 1 \\ 1/6 \\ -2/3 \end{bmatrix}$$

$$\vec{w}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 7/6 \\ 1 \\ 1/6 \\ -2/3 \end{bmatrix}}{\left\| \begin{bmatrix} 7/6 \\ 1 \\ 1/6 \\ -2/3 \end{bmatrix} \right\|^2} \begin{bmatrix} 7/6 \\ 1 \\ 1/6 \\ -2/3 \end{bmatrix}$$

$$- \frac{\begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}}{\left\| \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\|^2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

2. Normalize

$$\vec{q}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{\|\vec{w}_1\|} \vec{w}_1$$

$$\vec{q}_2 = \frac{1}{\|\vec{w}_2\|} \vec{w}_2$$

$$\vec{q}_3 = \frac{1}{\|\vec{w}_3\|} \vec{w}_3$$

Linear Transformations

A linear transformation T is a function from a vector space V to a vector space W such that

$$T(a\vec{v}+\vec{w}) = aT\vec{v} + T\vec{w}$$

$T(a\vec{v}) = aT\vec{v}$

Ex Is $T(x_1, x_2) = (2x_1 + x_2,\ x_2 - x_1)$ a linear transformation?

$T: \mathbb{R}^2 \to \mathbb{R}^2$

$a\vec{x}+\vec{y} = (ax_1+y_1,\ ax_2+y_2)$

$$\begin{aligned}
T(a\vec{x}+\vec{y}) &= T(ax_1+y_1,\ ax_2+y_2) \\
&= (2(ax_1+y_1) + ax_2+y_2,\\
&\qquad ax_2+y_2-(ax_1+y_1)) \\
&= (2ax_1 + ax_2,\ ax_2 - ax_1) \\
&\qquad + (2y_1+y_2,\ y_2-y_1) \\
&= a(2x_1+x_2,\ x_2-x_1) + (2y_1+y_2,\ y_2-y_1) \\
&= aT\vec{x} + T\vec{y}
\end{aligned}$$

Thus, T is a linear transformation

Ex Is there a linear transformation so that $T\alpha_i = \beta_i$ where

$$\alpha_1 = (2,1) \quad \beta_1 = (1,0)$$
$$\alpha_2 = (0,1) \quad \beta_2 = (2,2)$$
$$\alpha_3 = (2,0) \quad \beta_3 = (-1,-2).$$

Let $T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\quad T: \mathbb{R}^2 \to \mathbb{R}^2$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$2a + b = 1$$
$$2c + d = 0$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$b = 2 \Rightarrow a = -1/2$$
$$d = 2 \Rightarrow c = -1$$

$$T = \begin{bmatrix} -1/2 & 2 \\ -1 & 2 \end{bmatrix}$$

$$T\alpha_3 = \begin{bmatrix} -1/2 & 2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \beta_3 \checkmark$$

Positive definite matrix: Symmetric matrix with strictly positive eigenvalues

Test: $x^T A x > 0$ for all $\vec{x} \neq \vec{0}$.

arbitrary vector

$$\lambda = 0$$

$$A\vec{v} = \vec{0}$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \overset{?}{>} 0$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix}$$

$$2x_1^2 + x_1 x_2 + x_1 x_2 + 2x_2^2$$

$$2x_1^2 + 2x_1 x_2 + 2x_2^2$$

$$2(x_1^2 + x_1 x_2 + x_2^2) \overset{?}{>} 0$$

Similar matrices

$A = BCB^{-1}$, B is just some matrix

$$A = X \Lambda X^{-1} \leftarrow$$

Nullspace vs basis of Nullspace

$$\text{Nullspace of } A = \{ \vec{x} \mid A\vec{x} = \vec{0} \}$$

Basis of $N(A) = \{ \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \}$ so that

$\text{Span}(\vec{x}_1, \ldots, \vec{x}_k) = N(A)$

and $\vec{x}_1, \ldots, \vec{x}_k$ are linearly indep.

Ex

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

$$N(A) = \left\{ \begin{bmatrix} -2x_3 \\ x_2 \\ x_3 \end{bmatrix} \;\middle|\; x_2, x_3 \in \mathbb{R} \right\}$$

$$\begin{bmatrix} -2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$ form a basis of $N(A)$.

$$A\vec{x} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + x_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$