Review

$V = W \oplus W^{\perp}$

Definition

Let V be a vector space with an inner product $\langle .,. \rangle$.Let W be a subspace of V. We define the *orthogonal complement* of W in V to be the set

$$W^{\perp} = \{v \in V \mid \langle v, w \rangle = 0, \ \forall w \in W\}.$$

Proposition

Let V be a vector space with an inner product $\langle .,. \rangle$. Let W be a subspace of V. Then $W^{\perp}$ is a subspace of V and

$$W \cap W^{\perp} = \{0\}$$

Further, if V is finite-dimensional then $V = W + W^{\perp}$ and

$$\dim V = \dim W + \dim W^{\perp}$$

Orthogonal Decomposition Theorem

Theorem

Let V be a finite dimensional vector space with an inner product $\langle .,. \rangle$. Let W be a subspace of V. Then every vector $v \in V$ can be written uniquely in the form

$$v = w + \tilde{w}$$

$V \mapsto W$

where $w \in W$ and $\tilde{w} \in W^{\perp}$.

Definition

Let V, W, v, w and $\tilde{w}$ be as in the above theorem. We define the *orthogonal projection of v onto W* to be

$$\operatorname{proj}_W v = w$$

and the *component of v orthogonal to W* to be $\tilde{w} = v - \operatorname{proj}_W v$.

Proof: We already know that $V = W + W^\perp$, and that $W \cap W^\perp = \{0\}$.

Let $v \in V$.

Claim: The expression

$v = w + \tilde{w}$, where $w \in W, \tilde{w} \in W^\perp$

is unique, i.e.

if $v = w_1 + \tilde{w}_1 = w_2 + \tilde{w}_2$ where

$w_1, w_2 \in W$, and $\tilde{w}_1, \tilde{w}_2 \in W^\perp$

then $w_1 = \tilde{w}_1$, and $w_2 = \tilde{w}_2$.

Suppose $v = w_1 + \tilde{w}_1 = w_2 + \tilde{w}_2$, as above.

$\Rightarrow \quad w_1 - w_2 = \tilde{w}_2 - \tilde{w}_1$ — (1)

Now $w_1 - w_2 \in W$, $\tilde{w}_2 - \tilde{w}_1 \in W^\perp$ — (2)

From 1 & 2

$$w_1 - w_2 \in W \cap W^\perp = \{0\}$$

and $\tilde{w}_2 - \tilde{w}_1 \in W \cap W^\perp = \{0\}$

$$\therefore \quad w_1 = w_2, \quad \tilde{w}_1 = \tilde{w}_2 .$$

$\square$

Formula for Orthogonal Projection with respect to a Fixed Orthogonal Basis

Let V be a vector space of dimension n. Let $\langle ., . \rangle$ be an inner product defined on V.

Let W be a subspace of V.

Let $\{w_1, \ldots, w_m\}$ be an orthogonal basis of W and $\{w_{m+1}, \ldots, w_n\}$ be an orthogonal basis of $W^\perp$.

Let $v \in V$. Then

$$\operatorname{proj}_W v = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle v, w_m \rangle}{\langle w_m, w_m \rangle} w_m$$

and

$$v - \operatorname{proj}_W v = \frac{\langle v, w_{m+1} \rangle}{\langle w_{m+1}, w_{m+1} \rangle} w_{m+1} + \cdots + \frac{\langle v, w_n \rangle}{\langle w_n, w_n \rangle} w_n$$

Orthogonal Projection in $\mathbb{R}^n$

$\frac{y \cdot u_1}{u_1 \cdot u_1} = y \cdot u_1$

Theorem

If $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\,\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\,\mathbf{u}_2 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\,\mathbf{u}_p$$

If $U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p]$, then

$$\text{proj}_W \mathbf{y} = UU^T\mathbf{y}, \quad \forall \mathbf{y} \in \mathbb{R}^n$$

$A = [a_1 \; \dots \; a_n]$

$x = (x_1, \dots, x_n)$

$Ay = x_1 a_1 + \cdots + x_n a_n$

Proof:

$$\text{proj}_W \mathbf{y} = U \begin{bmatrix} (\mathbf{y} \cdot \mathbf{u}_1) \\ (\mathbf{y} \cdot \mathbf{u}_2) \\ \vdots \\ (\mathbf{y} \cdot \mathbf{u}_p) \end{bmatrix} = U \begin{bmatrix} \mathbf{u}_1^T \mathbf{y} \\ \mathbf{u}_2^T \mathbf{y} \\ \vdots \\ \mathbf{u}_p^T \mathbf{y} \end{bmatrix} = UU^T\mathbf{y}$$

$$U^T = \overset{p \times n}{\begin{bmatrix} u_1^T \\ \vdots \\ \vdots \\ u_p^T \end{bmatrix}} \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{bmatrix}$$

$$Uy = \begin{bmatrix} u_1^T \\ \vdots \\ \vdots \\ u_p^T \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} u_1^T \cdot y \\ \vdots \\ \vdots \\ u_p^T \cdot y \end{bmatrix}$$

Example

$U^T U = I$.

Compute $\text{proj}_W \mathbf{y}$, where

$$\mathbf{u}_1 = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$$

and

$$W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}.$$

$$UU^T = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

So $\operatorname{proj}_W \mathbf{y} = UU^T\mathbf{y} = \begin{bmatrix} -9 \\ \frac{7}{2} \\ \frac{7}{2} \end{bmatrix}$

QR Factorization

Definition

Let A be an $m \times n$ matrix having linearly independent columns. A *QR-factorization* of A is a matrix factorization of the form

$$A = QR$$

where Q is an $m \times n$ matrix having orthonormal columns and R is an upper triangular $n \times n$ matrix with strictly positive diagonal entries.

QR Algorithm

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the columns of A. Apply the Gram-Schmidt algorithm the the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ to obtain an orthogonal set $\mathbf{b}_1, \ldots, \mathbf{b}_n$, i.e.

$$\begin{aligned}\mathbf{b}_1 &= \mathbf{a}_1 \\ \mathbf{b}_k &= \mathbf{a}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{a}_k, \mathbf{b}_i \rangle}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle} \mathbf{b}_i, \quad \text{for } k = 2, \ldots, n.\end{aligned}$$

Normalize the vectors $\mathbf{b}_k$ to obtain an orthonormal set,

$$\mathbf{q}_k = \frac{\mathbf{b}_k}{\|\mathbf{b}_k\|}, \quad \text{for } k = 1, \ldots, n.$$

Thus we obtain the matrix

$$Q = [\mathbf{q}_1 \quad \dots \quad \mathbf{q}_n]$$

having orthonormal columns. Therefore $Q^T Q = I$. We put $R = Q^T A$, and this is gives us the required matrix factorization.

If time permits on Thursday, we will see a proof of why this works.

1 Why is R upper triangular with strictly positive diagonal entries?

2 Why is A = QR ?

Algebra of Linear Transformations; the space $L(V, W)$

Lemma

Let V and W be vector spaces. Let $\mathcal{U}$ be the set of all functions from V to W. Then $\mathcal{U}$ is a vector space under pointwise addition and pointwise scalar multiplication of functions.

Proof: Exercise. Verify all axioms. (HW)

Definition

Let V and W be vector spaces. The set of all linear transformations from V to W is called $L(V, W)$.

Proposition

$L(V, W)$ is a vector space under pointwise addition of functions and pointwise multiplication of functions by real numbers.

Proof: Let $T, S \in L(V,W)$

Claim: $T+S$ is a linear transformation.

Let $v, w \in V$.

$$(T+S)(v+w)$$

$$= \underline{T(v+w)} + \underline{S(v+w)}$$

$$= T(v) + T(w) + S(v) + S(w)$$

$$= \underbrace{T(v) + S(v)} + \underbrace{T(w) + S(w)}$$

$$= (T+S)(v) + (T+S)(w).$$

$\therefore$ Let $c \in \mathbb{R}$, $v \in V$.

$$(T+S)(cv) = T(cv) + S(cv)$$

$$= cT(v) + cS(v)$$

$$= c(T+S)(v).$$

$$\therefore \quad T+S \in L(V,W)$$

(V,W) If $T \in L(V,W)$ then $cT \in L(V,W)$ $\forall c \in \mathbb{R}$.

$T \in L(V, W) \longrightarrow [T]_{\beta, \gamma}$

Theorem

Let $\dim V = n$ *and* $\dim W = m$*. Then* $L(V, W)$ *is isomorphic to the vector space* $M_{m\times n}(\mathbb{R})$ *of* $m \times n$ *matrices.*

(In other words there exists a bijective linear transformation from $L(V, W)$ *to* $M_{m\times n}(\mathbb{R})$*).*

Fix a basis β for V
and a basis γ for W.

$$L(V, W) \xrightarrow{\alpha} M_{m\times n}(\mathbb{R})$$

$$T \longmapsto \text{matrix of } T$$

wrt β & $\mathcal{C}$.

$$\alpha(T) = [T]_{\beta,\mathcal{C}}.$$

We will show that $\alpha : L(V,W) \longrightarrow M_{m \times n}(\mathbb{R})$

sending $T \longmapsto [T]_{\beta,\varphi}$

is a linear transformation.

Let $T_1, T_2 \in L(V,W)$.

Claim: $\underbrace{[T_1 + T_2]_{\beta,\varphi}} = \underbrace{[T_1]_{\beta,\varphi} + [T_2]_{\beta\varphi}}$

Let $v \in V$.

$$[(T_1+T_2)(v)]_\varphi$$

$$= \left[T_1(v) + T_2(v) \right]_\varphi$$

$$= [T_1(v)]_\varphi + [T_2(v)]_\varphi$$

(because $x \mapsto [x]_\varphi$ is linear)

$$= \underline{[T_1]_{\beta,\mathcal{C}_\varphi}}\,\underline{[v]_\beta} + \underline{[T_2]_{\beta,\mathcal{C}_\varphi}}\,\underline{[v]_\beta}$$

$$= \left([T_1]_{\beta,\mathcal{C}_\varphi} + [T_2]_{\beta,\mathcal{C}_\varphi}\right)[v]_\beta .$$

$$\therefore \quad [T_1+T_2]_{\beta,\mathcal{C}_\varphi} = [T_1]_{\beta,\mathcal{C}_\varphi} + [T_2]_{\beta,\mathcal{C}_\varphi} .$$

$\therefore$ α respects vector addition in $L(V,W)$.

(HW): α also respects scalar multiplication.

i.e. $[cT]_{\beta,\mathcal{C}} = c[T]_{\beta,\mathcal{C}}$ $\forall c \in \mathbb{R}$, $\forall T \in L(V,W)$

Let us show that α is injective.

Let $T \in \ker \alpha$, i.e.

$\alpha(T) = 0$.

$\Rightarrow \quad [T(v)]_{\mathcal{C}} = \alpha(T)[v]_\beta = 0, \ \forall v \in V.$

$\Rightarrow \quad T(v) = 0, \ \forall v \in V. \Rightarrow T = 0.$

$\therefore$ α is injective.

Claim: α is surjective.

$$\begin{tikzcd}
V \arrow[r] \arrow[d, leftrightarrow] \arrow[d, bend right=60, "S_1"'] & W \arrow[d, leftrightarrow] \\
\mathbb{R}^n \arrow[r] & \mathbb{R}^m \arrow[u, bend right=60, "S_2"']
\end{tikzcd}$$

S_1 : coordinate map from $V \to \mathbb{R}^n$ sending

$$v \mapsto [v]_\beta$$

S_2 : coordinate map from

$$W \to \mathbb{R}^m$$

Sending

$$w \mapsto [w]_{e}.$$

Given a matrix

$A \in M_{m\times n}(\mathbb{R})$

define $T(v) = S_2^{-1}(A S_1(v))$.

This is composite of linear maps, therefore T linear.

$$[T(v)]_{\mathcal{C}}$$

$$= \left[\underline{S_2^{-1}}\left(AS_1(v)\right)\right]_{\mathcal{C}}$$

$$= AS_1(v)$$

$$= \underline{A}[v]_\beta \quad , \forall v \in V.$$

$$\therefore \quad A = [T]_{\beta, e_4}.$$

$$\begin{tikzcd} U \arrow[r, "S"] \arrow[rr, bend right, "T \circ S"'] & V \arrow[r, "T"] & W \end{tikzcd}$$

Theorem

Let U, V and W be finite dimensional vector spaces of dimensions m, n and k respectively, having bases $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ respectively.

Let $S \in L(U, V)$ and $T \in L(V, W)$. Then $T \circ S \in L(U, W)$ and

$$[T \circ S]_{\mathcal{A},\mathcal{C}} = [T]_{\mathcal{B},\mathcal{C}}[S]_{\mathcal{A},\mathcal{B}}$$

Proof: We first show $T \circ S \in L(U, W)$

Let $u_1, u_2 \in U$.

$$(T \circ S)(u_1 + u_2)$$

$$= T(S(u_1 + u_2))$$

$$= T(S(u_1) + S(u_2))$$

$$= T(S(u_1)) + T(S(u_2))$$

$= T \circ S(u_1) + (T \circ S)(u_2)$

$\therefore$ $T \circ S$ satisfies first condition for being a linear transformation.

Let $u \in U$, $c \in \mathbb{R}$.

$$T \circ S(cu)$$

$$= T(S(cu))$$

$$= T(c\,S(u))$$

$$= c\,T(S(u))$$

$$= c\,T \circ S(u).$$

$\therefore \ T \circ S \in \mathcal{L}(V, W)$.

~~let~~ Let $u \in U$.

$$[(T \circ S)(u)]_{\mathcal{C}}$$

$$= [T(S(u))]_{\mathcal{C}}$$

$$= [T]_{\mathcal{B},\mathcal{C}} \, [S(u)]_{\mathcal{B}}$$

$$= \left([T]_{\beta, \mathcal{C}} [S]_{A, \beta} \right) [u]_A$$

$$\therefore \quad [T \circ S]_{A, \mathcal{C}} = [T]_{\beta, \mathcal{C}} [S]_{A, \beta}.$$