21-241 MATRICES AND LINEAR TRANSFORMATIONS SUMMER 1 2012 COURSE NOTES JUNE 14

PAUL MCKENNEY

Definition. If V is a subspace of $\mathbb{C}^n$, then the *orthogonal subspace* of V is

$$V^\perp = \{w \in \mathbb{C}^n \mid \forall v \in V \ \ \langle w, v\rangle = 0\}$$

Fact 1. Let V be a subspace of $\mathbb{C}^n$.

(1) $V^\perp$ is a subspace of $\mathbb{C}^n$.
(2) $V \perp V^\perp$.
(3) $(V^\perp)^\perp = V$.

Theorem 1. *If V is a subspace of $\mathbb{C}^n$ and $x \in \mathbb{C}^n$, then $x - \mathbb{P}_V(x) \in V^\perp$.*

Proof. Let $\{v_1, \ldots, v_k\}$ be an orthonormal basis for V. Set $w = x - \mathbb{P}_V(x)$. We'll first show that w is orthogonal to each v_i in turn. To see this, recall that

$$\mathbb{P}_V(x) = \sum_{j=1}^{k} \langle x, v_j\rangle v_j$$

Then,

$$\begin{aligned}
\langle w, v_i\rangle &= \langle x - \mathbb{P}_V(x), v_i\rangle \\
&= \langle x, v_i\rangle - \left\langle \sum_{j=1}^{k} \langle x, v_j\rangle v_j, v_i \right\rangle \\
&= \langle x, v_i\rangle - \sum_{j=1}^{k} (\langle x, v_j\rangle) \langle v_j, v_i\rangle \\
&= \langle x, v_i\rangle - (0 + \cdots + \langle x, v_i\rangle \cdot 1 + \cdots + 0) \\
&= 0
\end{aligned}$$

Now let $v \in V$ be given. Since $\{v_1, \ldots, v_k\}$ is a basis for V, we may write

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

for some sequence of coefficients, $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$. Then,

$$\langle w, v\rangle = \langle w, \lambda_1 v_1 + \cdots + \lambda_k v_k\rangle = \overline{\lambda_1} \langle w, v_1\rangle + \cdots + \overline{\lambda_k} \langle w, v_k\rangle = 0 + \cdots + 0 = 0$$

□

Now we can describe the Gram-Schmidt process, which takes in a sequence of vectors $x_1, \dots, x_k$, and produces vectors $v_1, \dots, v_\ell$, and $u_1, \dots, u_\ell$, such that

(a) $\{v_1, \dots, v_\ell\}$ is an orthogonal basis for $\text{span}\{x_1, \dots, x_k\}$.
(b) $\{u_1, \dots, u_\ell\}$ is an orthonormal basis for $\text{span}\{x_1, \dots, x_k\}$.

The vectors v_i are computed recursively as follows;

(1) $v_1 = x_1$.
(2) $v_{i+1} = x_{i+1} - \mathbb{P}_{\text{span}\{v_1,\dots,v_i\}}(x_{i+1})$.

At each step we compute u_i by the formula

$$u_i = \frac{1}{\|v_i\|} v_i$$

if $v_i \neq 0$. If $v_i = 0$ then we discard x_i and compute v_i again with x_{i+1} in its place.

Note 1. Notice that, by construction (and the preceding theorem), v_{i+1} is in the orthogonal subspace of $\text{span}\{v_1, \dots, v_i\}$. In particular, v_{i+1} is orthogonal to each of $v_1, \dots, v_i$. Hence at the end we will have an orthogonal set of vectors. Moreover, at each step we've ensured that $v_1, \dots, v_i$ are pairwise orthogonal, so $u_1, \dots, u_i$ make up an orthonormal basis for their span. Hence by our definition of the projection operators $\mathbb{P}_S$, we can expand on our recursive definition of v_{i+1} above;

$$v_{i+1} = x_{i+1} - \mathbb{P}_{\text{span}\{v_1,\dots,v_i\}}(x_{i+1}) = x_{i+1} - \left(\sum_{j=1}^{i} \langle x_{i+1}, u_j \rangle \, u_j \right)$$

Note 2. If $x_1, \dots, x_k$ are already linearly independent, then it follows that $\ell = k$. Otherwise, it follows that $\ell < k$. (Why?)

Note 3. The case where one of the vectors v_i is zero only occurs when $x_i \in \text{span}\{x_1, \dots, x_{i-1}\}$, which itself only occurs when $x_1, \dots, x_k$ are linearly dependent.