4.3. Projection and Gram-Schmidt Process

我們曾經介紹 $\mathbb{R}^n$ 上的 projection. 簡單來說, 給定非零向量 $\mathbf{w}\in\mathbb{R}^n$, 我們定義 $\mathbf{v}\in\mathbb{R}^n$ 在 $\mathbf{w}$ 上的投影向量 $\mathbf{w}'$ 必須是在 $\mathbf{w}$ 所展成的空間中 (即 $\mathbf{w}'\in\mathrm{Span}(\mathbf{w})$), 而且 $\mathbf{v}-\mathbf{w}'$ 必須與 $\mathbf{w}$ 垂直, 也因此需與 $\mathrm{Span}(\mathbf{w})$ 上所有向量垂直. 依此, 我們將投影的定義推廣到一般對 inner product space V 的一個 subspace W 上的投影. 也就是說, 當 W 為 V 的 subspace, 對於 $\mathbf{v}\in V$, 我們定義 $\mathbf{v}$ 在 W 的 projection 為 W 上的一個向量 $\mathbf{w}'$ (即 $\mathbf{w}'\in W$), 滿足 $\mathbf{v}-\mathbf{w}$ 和 W 上所有向量垂直. 首先我們給以下的定義.

Definition 4.3.1. 假設 V 為 inner product space. 給定 W 為 V 的 subspace. 令

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

也就是說 $W^{\perp}$ 為 V 中和所有 W 中的向量垂直的向量所成的集合. 一般稱 $W^{\perp}$ 為 *orthogonal complement* of W.

有了 orthogonal complement 的定義我們就可以對上述的 projection 給了以下的定義.

Definition 4.3.2. 假設 V 為 inner product space. 給定 W 為 V 的 subspace. 對於 $\mathbf{v}\in V$, 若 $\mathbf{w}\in W$ 滿足 $\mathbf{v}-\mathbf{w}\in W^{\perp}$, 則稱 $\mathbf{w}$ 為 the *orthogonal projection* of $\mathbf{v}$ on W.

等一下我們將說明對於任意 $\mathbf{v}\in V$, the projection of $\mathbf{v}$ on W 一定存在, 而且唯一. 因此為了方便起見我們將此向量用 $\mathrm{Proj}_W(\mathbf{v})$ 表示. 我們將先證明唯一性, 有了唯一性, 將來我們要說明 $\mathbf{w}$ 就是 $\mathrm{Proj}_W(\mathbf{v})$, 就僅要檢查 $\mathbf{w}\in W$ 且滿足 $\mathbf{v}-\mathbf{w}\in W^{\perp}$ 就可以了.

我們先了解以下一些有關於 orthogonal complement 的性質. 假設 W 為 V 的 subspace, 並令 $\mathbf{v},\mathbf{v}'\in W^{\perp}$ 以及 $r,s\in\mathbb{R}$. 利用內積的性質, 對於任意 $\mathbf{w}\in W$, 我們有 $\langle r\mathbf{v}+s\mathbf{v}',\mathbf{w}\rangle=r\langle\mathbf{v},\mathbf{w}\rangle+s\langle\mathbf{v}',\mathbf{w}\rangle$. 再利用 $\mathbf{v},\mathbf{v}'\in W^{\perp}$, 即 $\langle\mathbf{v},\mathbf{w}\rangle=\langle\mathbf{v}',\mathbf{w}\rangle=0$, 得知 $\langle r\mathbf{v}+s\mathbf{v}',\mathbf{w}\rangle=0$. 此即表示對於任意 $\mathbf{v},\mathbf{v}'\in W^{\perp}$ 以及 $r,s\in\mathbb{R}$, 皆有 $r\mathbf{v}+s\mathbf{v}'\in W^{\perp}$, 得證 $W^{\perp}$ 為 V 的 subspace.

依定義要說明 $\mathbf{v}\in W^{\perp}$ 就得說明 $\langle\mathbf{v}.\mathbf{w}\rangle=0$, $\forall\mathbf{w}\in W$. 這個過程看似複雜, 因為要對所有 W 檢查. 不過由於 W 是 vector space, 我們僅要找到 W 的一組 basis, 然後對這組 basis 檢查即可, 因為我們有以下之結果.

Lemma 4.3.3. 假設 V 為 *inner product space* 且 W 為 V 的 *subspace.* 假設 $\mathbf{w}_1,\dots,\mathbf{w}_n$ 是 W 的一組 *basis,* 則 $\mathbf{v}\in W^{\perp}$ 若且唯若 $\langle\mathbf{v},\mathbf{w}_i\rangle=0$*,* $\forall i=1,\dots,n$*.*

Proof. 依定義若 $\mathbf{v}\in W^{\perp}$ 則對任意 $\mathbf{w}\in W$ 皆有 $\langle\mathbf{v},\mathbf{w}\rangle=0$, 所以當然 $\langle\mathbf{v},\mathbf{w}_i\rangle=0$, $\forall i=1,\dots,n$ 成立. 我們僅要證明若 $\langle\mathbf{v},\mathbf{w}_i\rangle=0$, $\forall i=1,\dots,n$ 則對任意 $\mathbf{w}\in W$ 皆有 $\langle\mathbf{v},\mathbf{w}\rangle=0$. 對任意 $\mathbf{w}\in W$, 因為 $\mathbf{w}_1,\dots,\mathbf{w}_n$ 是 W 的一組 basis, 故存在 $c_1,\dots,c_n\in\mathbb{R}$ 使得 $\mathbf{w}=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n$, 故得

$$\langle\mathbf{v},\mathbf{w}\rangle=\langle\mathbf{v},c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n\rangle=c_1\langle\mathbf{v},\mathbf{w}_1\rangle+\cdots+c_n\langle\mathbf{v},\mathbf{w}_n\rangle=0.$$

□

要注意 orthogonal complement 的 complement 不是指集合的補集. W 的 orthogonal complement $W^\perp$ 並不是 W 的補集. 甚至 $W\cap W^\perp$ 並不會是空集合. 這是因為 $W\cap W^\perp$ 也會是一個 subspace, 所以 $\mathbf{0}$ 一定在其中. 事實上我們有以下的結果.

Lemma 4.3.4. *假設 V 為 inner product space 且 W 為 V 的 subspace. 則 $W\cap W^\perp=\{\mathbf{0}\}$.*

Proof. 因 $W\cap W^\perp$ 為 subspace, 故知 $\mathbf{0}\in W\cap W^\perp$. 現假設 $\mathbf{w}\in W\cap W^\perp$. 由於 $\mathbf{w}\in W^\perp$, 對任意 $\mathbf{w}'\in W$ 皆有 $\langle\mathbf{w},\mathbf{w}'\rangle=0$. 而又 $\mathbf{w}\in W$, 故得 $\langle\mathbf{w},\mathbf{w}\rangle=0$. 因此由內積的性質知 $\mathbf{w}=\mathbf{0}$, 得證 $W\cap W^\perp=\{\mathbf{0}\}$. □

現在我們可以證明 projection 的唯一性.

Proposition 4.3.5. *假設 V 為 inner product space. 給定 W 為 V 的 subspace. 對於 $\mathbf{v}\in V$, $\mathbf{v}$ 對於 W 的 projection 是唯一的.*

Proof. 假設 $\mathbf{w},\mathbf{w}'\in W$ 皆為 $\mathbf{v}$ 在 W 的 projection. 也就是說 $\mathbf{w},\mathbf{w}'$ 皆滿足 $\mathbf{v}-\mathbf{w}\in W^\perp$ 以及 $\mathbf{v}-\mathbf{w}'\in W^\perp$. 由於 $W^\perp$ 為 subspace, 我們有 $(\mathbf{v}-\mathbf{w})-(\mathbf{v}-\mathbf{w}')\in W^\perp$, 亦即 $\mathbf{w}-\mathbf{w}'\in W^\perp$. 又因 W 為 subspace, 我們也知 $\mathbf{w}-\mathbf{w}'\in W$. 故由 Lemma 4.3.4 知 $\mathbf{w}-\mathbf{w}'=\mathbf{0}$, 即 $\mathbf{w}=\mathbf{w}'$, 證得唯一性. □

有了 Proposition 4.3.5 的唯一性, 以後我們要說明 $\mathbf{w}$ 是 $\mathbf{v}$ 在 W 的 projection, 就僅要檢查 $\mathbf{w}\in W$ 且滿足 $\mathbf{v}-\mathbf{w}\in W^\perp$ 就可以了.

接下來我們探討 projection 的存在性. 要考慮 $\mathbf{v}$ 對 W 的 projection, 我們必須找到 W 的一組 basis, 然後再看看這組 basis 所有可能的線性組合, 哪一個可以符合 projection 的要求. 這裡較麻煩的地方就是, 我們不知道要選取哪一個線性組合. 是否有可能找到一組特殊的 basis, 可以讓上述找到線性組合變簡單呢? 我們便是要回答這一個問題, 並探討如何找到這種特殊的 basis.

假設 W 為 V 的 subspace 且 $\mathbf{w}_1,\dots,\mathbf{w}_n$ 為其一組 basis 滿足 $\langle\mathbf{w}_i,\mathbf{w}_j\rangle=0,\forall i\neq j$. 對於任意 $\mathbf{w}\in W$, 因為 $\mathbf{w}_1,\dots,\mathbf{w}_n$ 為 W 的一組 basis, 存在 $c_1,\dots,c_n\in\mathbb{R}$ 使得 $\mathbf{w}=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n$. 一般來說我們都是利用解聯立方程組的方法找到 $c_1,\dots,c_n$, 不過這裡由於這些 $\mathbf{w}_i$ 之間兩兩互相垂直, 我們可以利用內積求出 c_i. 事實上對於任意 $i=1,\dots,n$, 考慮 $\langle\mathbf{w},\mathbf{w}_i\rangle$. 我們有

$$\langle\mathbf{w},\mathbf{w}_i\rangle=\langle c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n,\mathbf{w}_i\rangle=c_1\langle\mathbf{w}_1,\mathbf{w}_i\rangle+\cdots+c_n\langle\mathbf{w}_n,\mathbf{w}_i\rangle=c_i\langle\mathbf{w}_i,\mathbf{w}_i\rangle.$$

因為 $\mathbf{w}_i\neq\mathbf{0}$, 我們有 $\langle\mathbf{w}_i,\mathbf{w}_i\rangle=\|\mathbf{w}_i\|^2\neq 0$, 故得 $c_i=\langle\mathbf{w},\mathbf{w}_i\rangle/\langle\mathbf{w}_i,\mathbf{w}_i\rangle=\langle\mathbf{w},\mathbf{w}_i\rangle/\|\mathbf{w}_i\|^2$. 特別的, 若 $\|\mathbf{w}_i\|=1$, 則 $c_i=\langle\mathbf{w},\mathbf{w}_i\rangle$. 我們有以下的定理.

Proposition 4.3.6. *假設 V 為 inner product space, W 為 V 的 subspace 且 $\mathbf{w}_1,\dots,\mathbf{w}_n$ 為 W 的一組 basis 滿足 $\langle\mathbf{w}_i,\mathbf{w}_j\rangle=0,\forall i\neq j$. 則對於任意 $\mathbf{w}\in W$, 我們有*

$$\mathbf{w}=\frac{\langle\mathbf{w},\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1+\cdots+\frac{\langle\mathbf{w},\mathbf{w}_i\rangle}{\|\mathbf{w}_i\|^2}\mathbf{w}_i+\cdots+\frac{\langle\mathbf{w},\mathbf{w}_n\rangle}{\|\mathbf{w}_n\|^2}\mathbf{w}_n.$$

由於這種兩兩互相垂直的 basis, 對於寫下一個向量的 linear combination 相當的方便, 我們有以下的定義.

Definition 4.3.7. 假設 V 為 inner product space 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 basis 滿足對於任意 $i\neq j$ 皆有 $\langle \mathbf{v}_i,\mathbf{v}_j\rangle=0$. 則稱 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 *orthogonal basis.* 若又要求 $\langle \mathbf{v}_i,\mathbf{v}_i\rangle=1$ (即 $\|\mathbf{v}_i\|=1$), $\forall i=1,\ldots,n$, 則稱 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 V 的一組 *orthonormal basis.*

要注意, 若 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 V 的一組 orthogonal basis, 對於所有 $i=1,\ldots,n$, 令 $\mathbf{u}_i=\mathbf{v}_i/\|\mathbf{v}_i\|$, 則 $\mathbf{u}_1,\ldots,\mathbf{u}_n$ 就會是 V 的一組 orthonormal basis. 這是因為

$$\langle \mathbf{u}_i,\mathbf{u}_j\rangle=\langle\frac{1}{\|\mathbf{v}_i\|}\mathbf{v}_i,\frac{1}{\|\mathbf{v}_j\|}\mathbf{v}_j\rangle=\frac{1}{\|\mathbf{v}_i\|}\frac{1}{\|\mathbf{v}_j\|}\langle \mathbf{v}_i,\mathbf{v}_j\rangle.$$

因此當 $i\neq j$, 我們有 $\langle \mathbf{u}_i,\mathbf{u}_j\rangle=\langle \mathbf{v}_i,\mathbf{v}_j\rangle=0$; 而當 $i=j$, 我們有 $\langle \mathbf{u}_i,\mathbf{u}_i\rangle=\langle \mathbf{v}_i,\mathbf{v}_i\rangle/\|\mathbf{v}_i\|^2=1$.

由前面已知, 若能找到 W 的一組 orthogonal basis $\mathbf{w}_1,\ldots,\mathbf{w}_n$, 則我們可以很容易的將任意 W 中的向量寫成 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 的線性組合. 這似乎克服了前面所述較複雜的部份. 事實上確實如此, 若能找到 W 的一組 orthogonal basis $\mathbf{w}_1,\ldots,\mathbf{w}_n$, 則我們便可以輕易地得到 $\mathbf{v}$ 在 W 的 projection 了. 這是因為若 $\mathbf{w}=\mathrm{Proj}_W(\mathbf{v})$ 且 $\mathbf{w}=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n$. 此時由於 $\mathbf{v}-\mathbf{w}\in W^{\perp}$, 對於所有 $i=1,\ldots,n$, 我們有 $\langle \mathbf{v},\mathbf{w}_i\rangle-\langle \mathbf{w},\mathbf{w}_i\rangle=\langle \mathbf{v}-\mathbf{w},\mathbf{w}_i\rangle=0$. 因此得 $\langle \mathbf{v},\mathbf{w}_i\rangle=\langle \mathbf{w},\mathbf{w}_i\rangle=c_i\langle \mathbf{w}_i,\mathbf{w}_i\rangle$, 亦即 $c_i=\langle \mathbf{v},\mathbf{w}_i\rangle/\langle \mathbf{w}_i,\mathbf{w}_i\rangle$. 所以我們得到一個很簡捷求 projection 的方法.

Theorem 4.3.8. 假設 V 為 *inner product space,* W 為 V 的 *subspace* 且 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 為 W 的一組 *orthogonal basis.* 若 $\mathbf{v}\in V$, 則 $\mathbf{v}$ 在 W 的 *projection* 為

$$\mathrm{Proj}_W(\mathbf{v})=\frac{\langle \mathbf{v},\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1+\cdots+\frac{\langle \mathbf{v},\mathbf{w}_n\rangle}{\|\mathbf{w}_n\|^2}\mathbf{w}_n.$$

特別的當 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 為 W 的一組 *orthonormal basis,* 則

$$\mathrm{Proj}_W(\mathbf{v})=\langle \mathbf{v},\mathbf{w}_1\rangle\mathbf{w}_1+\cdots+\langle \mathbf{v},\mathbf{w}_n\rangle\mathbf{w}_n.$$

Proof. 令 $\mathbf{w}=c_1\mathbf{w}_1+\cdots+c_n\mathbf{w}_n$, 其中 $c_i=\langle \mathbf{v},\mathbf{w}_i\rangle/\langle \mathbf{w}_i,\mathbf{w}_i\rangle$, $\forall i=1,\ldots,n$. 此時 $\mathbf{w}\in W$, 且對任意 $i=1,\ldots,n$ 皆有

$$\langle \mathbf{v}-\mathbf{w},\mathbf{w}_i\rangle=\langle \mathbf{v},\mathbf{w}_i\rangle-\langle \mathbf{w},\mathbf{w}_i\rangle=\langle \mathbf{v},\mathbf{w}_i\rangle-\sum_{j=1}^{n}c_j\langle \mathbf{w}_j,\mathbf{w}_i\rangle=\langle \mathbf{v},\mathbf{w}_i\rangle-c_i\langle \mathbf{w}_i,\mathbf{w}_i\rangle=0.$$

故由 Lemma 4.3.3 知 $\mathbf{v}-\mathbf{w}\in W^{\perp}$. 因此由 projection 的唯一性 (Proposition 4.3.5) 知 $\mathbf{w}=\mathrm{Proj}_W(\mathbf{v})$. □

我們已經知道只要找到 W 的一組 orthogonal basis, 就可以輕易求得 $\mathbf{v}$ 在 W 的 projection. 對一般 inner product space, 我們將介紹一個方法找到它的一組 orthogonal basis, 也因此證明了對於一般的 inner product space 一定存在 orthogonal basis (以及 orthonormal basis). 這個方法就是所謂的 *Gram-Schmidt process.*

給定 V 的一個 nonzero subspace W, 且假設 $\dim(W)=n$. 首先我們說明若 $\mathbf{w}_1,\ldots,\mathbf{w}_k\in W$ 為非零向量且滿足 $\langle \mathbf{w}_i,\mathbf{w}_j\rangle=0$, $\forall i\neq j$, 則 $\mathbf{w}_1,\ldots,\mathbf{w}_k$ 為 linearly independent. 這是因為若不是 linearly independent 表示存在 $c_1,\ldots,c_k\in\mathbb{R}$ 不全為 0 使得 $c_1\mathbf{w}_1+\cdots+c_k\mathbf{w}_k=\mathbf{0}$. 然而對任意 $i=1,\ldots,k$ 由於 $0=\langle c_1\mathbf{w}_1+\cdots+c_k\mathbf{w}_k,\mathbf{w}_i\rangle=c_i\|\mathbf{w}_i\|^2$. 也因此由 $\|\mathbf{w}_i\|\neq 0$, 得證 $c_i=0$. 此和 $c_1,\ldots,c_k$ 不全為 0 的假設相矛盾, 故知 $\mathbf{w}_1,\ldots,\mathbf{w}_k$ 為 linearly independent. 因此要找

到 W 的一組 orthogonal basis, 我們只要在 W 中找到 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 滿足 $\langle \mathbf{w}_i,\mathbf{w}_j\rangle = 0, \forall i\neq j$ 即可, 因為它們是 linearly independent 且 $\dim(W)=n$, 故知它們是 W 的一組 basis. 接下來我們要說明在 W 中如何找到這樣的一組 nonzero vectors.

首先因 $W\neq\{\mathbf{0}\}$, 故可在 W 中取一 nonzero vector $\mathbf{v}_1$. 為了方便起見我們令 $\mathbf{w}_1=\mathbf{v}_1$ 且 $W_1=\mathrm{Span}(\mathbf{v}_1)=\mathrm{Span}(\mathbf{w}_1)$. 若 $\dim(W)=1$, 則 $W=W_1$ 故 $\mathbf{w}_1$ 就是 W 的一個 orthogonal basis. 而若 $\dim(W)>1$, 則因 $W_1\subsetneq W$, 我們可以找到 nonzero vector $\mathbf{v}_2\in W$ 且 $\mathbf{v}_2\not\in W_1$. 現在我們要利用 $\mathbf{v}_2$, 找到 $\mathbf{w}_2\in W$ 滿足 $\mathbf{w}_2\neq\mathbf{0}$ 且 $\langle\mathbf{w}_1,\mathbf{w}_2\rangle=0$. 很自然的, 我們會考慮 $\mathbf{w}_2=\mathbf{v}_2-\mathrm{Proj}_{W_1}(\mathbf{v}_2)$, 因為此時 $\mathbf{w}_2\in W_1^{\perp}$, 而 $W_1=\mathrm{Span}(\mathbf{w}_1)$, 故當然有 $\langle\mathbf{w}_1,\mathbf{w}_2\rangle=0$. 我們也要說明 $\mathbf{w}_2\neq\mathbf{0}$. 這是因為若 $\mathbf{w}_2=\mathbf{0}$, 會得到 $\mathbf{v}_2=\mathrm{Proj}_{W_1}(\mathbf{v}_2)\in W_1$, 此與當初 $\mathbf{v}_2\not\in W_1$ 的假設相矛盾. 另一方面因為 $W=\mathrm{Span}(\mathbf{w}_1)$, 利用 Proposition 4.1.9 我們知 $\mathrm{Proj}_{W_1}(\mathbf{v}_2)=(\langle\mathbf{v}_2,\mathbf{w}_1\rangle/\|\mathbf{w}_1\|^2)\mathbf{w}_1$, 所以我們知

$$\mathbf{w}_2=\mathbf{v}_2-\frac{\langle\mathbf{v}_2,\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1.$$

另外要注意的是 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)=\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$, 這是因為依 $\mathbf{w}_1,\mathbf{w}_2$ 的選取, 我們有 $\mathbf{w}_1,\mathbf{w}_2\in\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$, 因此 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)\subseteq\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$. 然而因為 $\mathbf{v}_1,\mathbf{v}_2$ 為 linearly independent 且 $\mathbf{w}_1,\mathbf{w}_2$ 為 linearly independent, 故由 $\dim(\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2))=\dim(\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2))=2$ 得證 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)=\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$. 為了方便起見, 我們令 $W_2=\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)=\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2)$. 現若 $\dim(W)=2$, 則 $W=W_2$, 故 $\mathbf{w}_1,\mathbf{w}_2$ 為 W 的一組 orthogonal basis. 而若 $\dim(W)>2$, 則因 $W_2\subsetneq W$, 我們可以找到 nonzero vector $\mathbf{v}_3\in W$ 且 $\mathbf{v}_3\not\in W_2$. 現在我們要利用 $\mathbf{v}_3$, 找到 $\mathbf{w}_3\in W$ 滿足 $\mathbf{w}_3\neq\mathbf{0}$ 且 $\langle\mathbf{w}_1,\mathbf{w}_3\rangle=\langle\mathbf{w}_2,\mathbf{w}_3\rangle=0$. 同前, 我們考慮 $\mathbf{w}_3=\mathbf{v}_3-\mathrm{Proj}_{W_2}(\mathbf{v}_3)$, 因為此時 $\mathbf{w}_3\in W_2^{\perp}$, 而 $W_2=\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2)$, 故當然有 $\langle\mathbf{w}_1,\mathbf{w}_3\rangle=\langle\mathbf{w}_2,\mathbf{w}_3\rangle=\mathbf{0}$. 另外因 $\mathbf{v}_3\not\in W_2$, 同前面的理由我們有 $\mathbf{w}_3\neq\mathbf{0}$. 另一方面因為 $\mathbf{w}_1,\mathbf{w}_2$ 為 W_2 的 orthogonal basis, 利用 Theorem 4.3.8 我們得

$$\mathbf{w}_3=\mathbf{v}_3-\mathrm{Proj}_{W_2}(\mathbf{v}_3)=\mathbf{v}_3-\frac{\langle\mathbf{v}_3,\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1-\frac{\langle\mathbf{v}_3,\mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2.$$

最後和前面同樣的理由, 我們有 $\mathrm{Span}(\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3)=\mathrm{Span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)$. 這樣一直下去, 我們可以得到 $W_k=\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k)=\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_k)$ 且 $\mathbf{w}_1,\ldots,\mathbf{w}_k$ 是 W_k 的一組 orthogonal basis. 現若 $k=\dim(W)=n$, 則得 $W_k=W$, 所以 $\mathbf{w}_1,\ldots,\mathbf{w}_k$ 是 W 的一組 orthogonal basis. 而若 $k<n$, 則存在 $\mathbf{v}_{k+1}\in W$ 且 $\mathbf{v}_{k+1}\not\in W_k$. 故令

$$\mathbf{w}_{k+1}=\mathbf{v}_{k+1}-\mathrm{Proj}_{W_k}(\mathbf{v}_{k+1})=\mathbf{v}_{k+1}-\frac{\langle\mathbf{v}_{k+1},\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1-\frac{\langle\mathbf{v}_{k+1},\mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2-\cdots-\frac{\langle\mathbf{v}_{k+1},\mathbf{w}_k\rangle}{\|\mathbf{w}_k\|^2}\mathbf{w}_k,$$

則得 $\mathbf{w}_{k+1}\neq\mathbf{0}$ 且 $\langle\mathbf{w}_{k+1},\mathbf{w}_i\rangle=0, \forall i=1,\ldots,k$. 另外因 $\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{w}_{k+1}\in\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_k,\mathbf{v}_{k+1})$, 同上可得 $\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{w}_{k+1})=\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_k,\mathbf{v}_{k+1})$, 故令 $W_{k+1}=\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{w}_{k+1})=\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_k,\mathbf{v}_{k+1})$, 我們有 $\mathbf{w}_1,\ldots,\mathbf{w}_k,\mathbf{w}_{k+1}$ 為 W_{k+1} 的一組 orthogonal basis. 這樣一直下去直到得到 $W_n=W$, 這樣 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 就是 W 的一組 orthogonal basis.

上述 Gram-Schmidt process 中 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的選取事實上和我們過去找 vector space 的 basis 方法是一樣的. 差別就是我們要將 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 這組 basis 修改成 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 這一組 orthogonal basis. 因此如果一開始已給定 W 的一組 basis $\mathbf{v}_1\ldots,\mathbf{v}_n$, 我們可以將之直接套用, 因此有以下的結果.

Theorem 4.3.9 (Gram-Schmidt Process)**.** 假設 V 為 *inner product space,* W 為 V 的 *subspace* 且 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為 W 的一組 *basis.* 令

$$\mathbf{w}_1=\mathbf{v}_1,\quad \mathbf{w}_2=\mathbf{v}_2-\frac{\langle\mathbf{v}_2,\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1,\ldots$$

這樣一直下去, 即對於 $i=1,\ldots,n-1$ 令

$$\mathbf{w}_{i+1}=\mathbf{v}_{i+1}-\frac{\langle\mathbf{v}_{i+1},\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1-\frac{\langle\mathbf{v}_{i+1},\mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2-\cdots-\frac{\langle\mathbf{v}_{i+1},\mathbf{w}_i\rangle}{\|\mathbf{w}_i\|^2}\mathbf{w}_i,$$

則 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 為 W 的一組 *orthogonal basis.* 而且

$$\mathrm{Span}(\mathbf{w}_1,\ldots,\mathbf{w}_i)=\mathrm{Span}(\mathbf{v}_1,\ldots,\mathbf{v}_i),\quad \forall i=1,\ldots,n.$$

Gram-Schmidt process 確保了 orthogonal basis 的存在性, 而當 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 為其 orthogonal basis 時, 我們可以除去其長度得到 $(1/\|\mathbf{v}_1\|)\mathbf{v}_1,\ldots,(1/\|\mathbf{v}_n\|)\mathbf{v}_n$ 這一組 orthonormal basis. 利用 orthogonal basis 的存在性, 我們也就得到了 orthogonal projection 的存在性了. 也就是說當 V 為 inner product space 且 W 為其 subspace, 我們就可以利用 Gram-Schmidt process 找到 W 的一組 orthogonal basis, 然後利用 Theorem 4.3.8, 得到任意 V 中的向量 $\mathbf{v}$ 在 W 上的 orthogonal projection 了.

Example 4.3.10. 我們要用 orthogonal basis 來處理 orthogonal projection 的問題. 我們在 $\mathbb{R}^4$ 使用 dot product, 考慮 $\mathbf{v}=\begin{bmatrix}2\\0\\1\\4\end{bmatrix}$ 在 $W=\mathrm{Span}(\begin{bmatrix}1\\2\\2\\1\end{bmatrix},\begin{bmatrix}3\\4\\2\\3\end{bmatrix})$ 的 orthogonal projection. 首先找 W 的一組 orthogonal basis. 令

$$\mathbf{w}_1=\begin{bmatrix}1\\2\\2\\1\end{bmatrix},\quad \mathbf{w}_2=\begin{bmatrix}3\\4\\2\\3\end{bmatrix}-\frac{18}{10}\begin{bmatrix}1\\2\\2\\1\end{bmatrix}=\frac{2}{5}\begin{bmatrix}3\\1\\-4\\3\end{bmatrix}.$$

此時 $\mathbf{w}_1,\mathbf{w}_2$ 為 W 的一組 orthogonal basis, 故利用 Theorem 4.3.8 得

$$\mathrm{Proj}_W(\mathbf{v})=\frac{\langle\mathbf{v},\mathbf{w}_1\rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1+\frac{\langle\mathbf{v},\mathbf{w}_2\rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2=\frac{8}{10}\begin{bmatrix}1\\2\\2\\1\end{bmatrix}+\frac{28/5}{28/5}\times\frac{2}{5}\begin{bmatrix}3\\1\\-4\\3\end{bmatrix}=\begin{bmatrix}2\\2\\0\\2\end{bmatrix}.$$

雖然 Theorem 4.3.9 的敘述是找到 V 的 subspace W 的 orthogonal basis, 不過因 W 是 V 中任意的 subspace, 所以當 W 為 V 時, 我們也就找到 V 的 orthogonal basis 了. 所以對任意 finite dimensional inner product space, Gram-Schmidt process 都能幫我們找到 orthogonal basis. 注意這裡需要有限維的假設, 因為整個過程我們是一個一個置換這些向量, 所以有限多個向量才可全部置換完成. Theorem 4.3.9 的敘述牽涉到 V 的 subspace W 主要用意是, 我們不只可找到 W 的 orthogonal basis, 也可繼續這個 process, 而將 W 的 orthogonal basis 擴大成 V 的 orthogonal basis. 這是因為若 $W\neq V$, 當找到 $\mathbf{w}_1,\ldots,\mathbf{w}_n$ 為 W 的 orthogonal basis 後我們可以繼續考慮 $\mathbf{v}_{n+1}\in V$ 但 $\mathbf{v}_{n+1}\notin W$, 然後利用 Gram-Schmidt process 得到

$\mathbf{w}_{n+1} = \mathbf{v}_{n+1} - \mathrm{Proj}_W(\mathbf{v}_{n+1}) \in W^{\perp}$. 這樣一直下去直到得到 V 的一組 orthogonal basis 為止. 也因此我們得到以下之結果.

Corollary 4.3.11. 假設 V 為 *inner product space,* W 為 V 的 *subspace.* 若 $\dim(V) = m$ 且 $\dim(W) = n$, 則存在 V 的一組 *orthogonal basis* $\mathbf{v}_1, \ldots, \mathbf{v}_n, \ldots, \mathbf{v}_m$, 其中 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 W 的 *orthogonal basis.*

Example 4.3.12. 考慮 $\mathbb{R}^4$ 中以 dot product 所形成的 inner product space. 令

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix},$$

我們要求 $W = \mathrm{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ 的一組 orthogonal basis, 並將之擴大成 $\mathbb{R}^4$ 的一組 orthogonal basis. 首先令 $\mathbf{w}_1 = \mathbf{v}_1$, 得

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix} - \frac{4}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 0 \end{bmatrix}.$$

最後得

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix} - \frac{8}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-4}{8}\begin{bmatrix} 2 \\ 0 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}.$$

由於 $\dim(V) = 3$ 且 $\dim(\mathbb{R}^4) = 4$, 我們需要再找到一個向量以形成 $\mathbb{R}^4$ 的 basis. 考慮 $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, 我們可以檢查 $\mathbf{v}_4 \notin W$ (或直接套用 Gram-Schmidt process, 若 $\mathbf{v}_4 \in W$, 會得到 $\mathbf{v}_4 - \mathrm{Proj}_W(\mathbf{v}_4) = \mathbf{0}$. 若真如此就再換一個向量). 得

$$\mathbf{w}_4 = \mathbf{v}_4 - \frac{\langle \mathbf{v}_4, \mathbf{w}_1 \rangle}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\langle \mathbf{v}_4, \mathbf{w}_2 \rangle}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \frac{\langle \mathbf{v}_4, \mathbf{w}_3 \rangle}{\|\mathbf{w}_3\|^2}\mathbf{w}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

此時 $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ 就是 $\mathbb{R}^4$ 的一組 orthogonal basis, 其中 $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ 是 W 的 orthogonal basis.

對於每一個 $\mathbf{w}_i$, 我們可以除以其長度 $\|\mathbf{w}_i\|$, 得到一組 orthonormal basis

$$\mathbf{u}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \mathbf{u}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_4 = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

接下來我們看幾個有關 orthogonal basis 的應用. 由於利用 orthonormal basis 會比較方便 (省去除掉長度的麻煩), 所以以下都用 orthonormal basis 處理. 首先當 W 是 inner product space V 的 subspace, 我們知道 $W^\perp$ 也是 V 的 subspace. 很自然的我們會想要知道 $\dim(W^\perp)$ 和 $\dim(W)$ 的關係. 當我們利用 Corollary 4.3.11 找到 $\mathbf{u}_1,\dots,\mathbf{u}_n,\dots,\mathbf{u}_m$ 是 V 的一組 orthonormal basis, 其中 $\mathbf{u}_1,\dots,\mathbf{u}_n$ 為 W 的 orthonormal basis, 我們就可以將任意 $\mathbf{v}\in W^\perp$ 寫成 $\mathbf{v}=c_1\mathbf{u}_1+\cdots+c_n\mathbf{u}_n+c_{n+1}\mathbf{u}_{n+1}+\cdots+c_m\mathbf{u}_m$. 其中 $c_i=\langle\mathbf{v},\mathbf{u}_i\rangle$ (Proposition 4.3.6 套用 $W=V$ 的情形). 由於 $\mathbf{v}\in W^\perp$, 我們有 $\langle\mathbf{v},\mathbf{u}_i\rangle=0$, $\forall i=1,\dots,n$ (因為這些 $\mathbf{u}_i\in W$). 因此得 $\mathbf{v}=c_{n+1}\mathbf{u}_{n+1}+\cdots+c_m\mathbf{u}_m\in\mathrm{Span}(\mathbf{u}_{n+1},\dots,\mathbf{u}_m)$. 反之若 $\mathbf{v}=c_{n+1}\mathbf{u}_{n+1}+\cdots+c_m\mathbf{u}_m\in\mathrm{Span}(\mathbf{u}_{n+1},\dots,\mathbf{u}_m)$, 由於當 $i\neq j$ 時 $\langle\mathbf{u}_i,\mathbf{u}_j\rangle=0$, 故當 $i=1,\dots,n$ 時

$$\langle\mathbf{v},\mathbf{u}_i\rangle=\sum_{j=n+1}^{m}c_j\langle\mathbf{u}_j,\mathbf{u}_i\rangle=0.$$

因此由 $\mathbf{u}_1,\dots,\mathbf{u}_n$ 為 W 的 basis 以及 Lemma 4.3.3 得 $\mathbf{v}\in W^\perp$, 也因此我們證明了 $W^\perp=\mathrm{Span}(\mathbf{u}_{n+1},\dots,\mathbf{u}_m)$. 又因為 $\mathbf{u}_{n+1},\dots,\mathbf{u}_m$ 為 linearly independent, 故知 $\mathbf{u}_{n+1},\dots,\mathbf{u}_m$ 為 $W^\perp$ 的一組 basis (事實上也是 orthonormal basis).

Proposition 4.3.13. 假設 V 為 *inner product space,* W 為 V 的 *subspace* 且設 $\dim(V)=m$, $\dim(W)=n$. 若 $\mathbf{v}_1,\dots,\mathbf{v}_n,\dots,\mathbf{v}_m$ 為 V 的一組 *orthogonal basis* 且其中 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 是 W 的 *orthogonal basis,* 則 $\mathbf{v}_{n+1},\dots,\mathbf{v}_m$ 為 $W^\perp$ 的 *orthogonal basis.* 特別的, 我們有

$$\dim(W^\perp)=\dim(V)-\dim(W).$$

Question 4.8. 在 *Example 4.3.12* 中, 試找到 $W^\perp$ 的一組 *basis.*

在補集的概念中, 我們知道一個集合的補集再取補集, 會是該集合本身. orthogonal complement 會不會也有同樣的情形呢? 也就是說當 W 是 inner product space V 的 subspace, 會不會有 $(W^\perp)^\perp=W$ 的情形發生? 一般要說明 $(W^\perp)^\perp=W$, 我們需證明 $W\subseteq(W^\perp)^\perp$ 以及 $(W^\perp)^\perp\subseteq W$. 證明 $W\subseteq(W^\perp)^\perp$ 這部分是簡單的, 因為 $(W^\perp)^\perp$ 依定義是所有和 $W^\perp$ 垂直的向量所成的集合, 所以當 $\mathbf{w}\in W$, 我們要說明 $\mathbf{w}\in(W^\perp)^\perp$ 僅要說明 $\langle\mathbf{w},\mathbf{v}\rangle=0$, $\forall\mathbf{v}\in W^\perp$ 即可. 然而任意 $\mathbf{v}\in W^\perp$ 依定義皆會和所有 W 中的向量垂直, 故由 $\mathbf{w}\in W$, 我們自然有 $\langle\mathbf{w},\mathbf{v}\rangle=0$, $\forall\mathbf{v}\in W^\perp$. 至於 $(W^\perp)^\perp\subseteq W$, 很不幸的它並不一定會成立. 事實上在 V 為無限維時可以找到反例. 由於本課程並不涉及無限維的向量空間, 這裡就略去不談. 不過在 V 為 finite dimensional inner product space, $(W^\perp)^\perp\subseteq W$ 就會成立. 只是它的證明是無法像前面 $W\subseteq(W^\perp)^\perp$ 的情況用集合元素方式推導 (否則就不會有在無限維時不成立的情況發生). 既然是有限維, 我們可以用為維度處理. 回顧一下當我們有 W' 為 W 的 subspace, 且知 $\dim(W')=\dim(W)$, 則可得 $W'=W$. 所以既然我們已知 $W\subseteq(W^\perp)^\perp$, 只要說明 $\dim(W)=\dim((W^\perp)^\perp)$, 就可得證 $(W^\perp)^\perp=W$.

Corollary 4.3.14. 假設 V 為 *finite dimensional inner product space* 且 W 為 V 的 *subspace.* 則 $(W^\perp)^\perp=W$.

Proof. 前面已證得 $W\subseteq(W^\perp)^\perp$, 所以現在僅要說明 $\dim(W)=\dim((W^\perp)^\perp)$, 就可得證 $(W^\perp)^\perp=W$. 然而由 Proposition 4.3.13, 我們知 $\dim((W^\perp)^\perp)=\dim(V)-\dim(W^\perp)$, 再由

$\dim(W^{\perp}) = \dim(V) - \dim(W)$, 得

$$\dim((W^{\perp})^{\perp}) = \dim(V) - (\dim(V) - \dim(W)) = \dim(W).$$

$\square$

在 $\mathbb{R}^2$, $\mathbb{R}^3$ 中的投影概念中, 還有一個重要的觀點就是一個點在直線 (或平面上) 的投影點就是這個線上 (或平面上) 距離該點最近的點. 這個概念對我們推廣到 inner product space 後的 orthogonal projection 也是對的. 我們有以下的性質.

Proposition 4.3.15. 假設 V 為 *inner product space* 且 W 為 V 的 *subspace.* 若 $\mathbf{w} = \mathrm{Proj}_W(\mathbf{v})$ 為 $\mathbf{v}$ 在 W 的 *orthogonal projection,* 則對於任意 $\mathbf{w}' \in W$ 且 $\mathbf{w}' \neq \mathbf{w}$, 皆有 $\|\mathbf{v}-\mathbf{w}'\| > \|\mathbf{v}-\mathbf{w}\|$.

Proof. 考慮 $\mathbf{v}-\mathbf{w}' = \mathbf{v}-\mathbf{w}+\mathbf{w}-\mathbf{w}'$. 因 $\mathbf{w} = \mathrm{Proj}_W(\mathbf{v})$, 故知 $\mathbf{v}-\mathbf{w} \in W^{\perp}$. 又因 W 為 vector space, 我們有 $\mathbf{w}-\mathbf{w}' \in W$. 故得

$$\|\mathbf{v}-\mathbf{w}'\|^2 = \langle \mathbf{v}-\mathbf{w}', \mathbf{v}-\mathbf{w}' \rangle = \langle \mathbf{v}-\mathbf{w}+\mathbf{w}-\mathbf{w}', \mathbf{v}-\mathbf{w}+\mathbf{w}-\mathbf{w}' \rangle = \langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w} \rangle + \langle \mathbf{w}-\mathbf{w}', \mathbf{w}-\mathbf{w}' \rangle.$$

亦即 $\|\mathbf{v}-\mathbf{w}'\|^2 = \|\mathbf{v}-\mathbf{w}\|^2 + \|\mathbf{w}-\mathbf{w}'\|^2$. 又因 $\mathbf{w} \neq \mathbf{w}'$ 我們有 $\|\mathbf{w}-\mathbf{w}'\| > 0$. 得證 $\|\mathbf{v}-\mathbf{w}'\|^2 > \|\mathbf{v}-\mathbf{w}\|^2$, 即 $\|\mathbf{v}-\mathbf{w}'\| > \|\mathbf{v}-\mathbf{w}\|$. $\square$