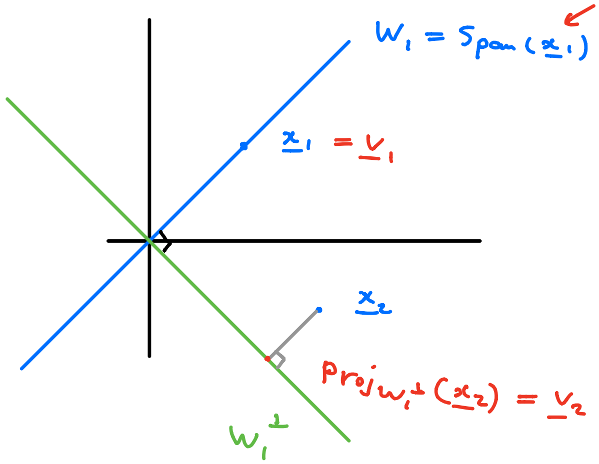
The Gram-Schmidt Process

Aim: Given $W \subset \mathbb{R}^n$ a subspace find $\{\underline{v}_1, \dots, \underline{v}_p\}$ an orthogonal basis.

Gram-Schmidt Process: Algorithm to construct $\{\underline{v}_1, \dots, \underline{v}_p\}$ starting from a basis $\{\underline{x}_1, \dots, \underline{x}_p\} \subset W$.

Example $W = \mathbb{R}^2$, $W = \text{Span}\left(\underset{\underset{x_1}{\shortparallel}}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}, \underset{\underset{x_2}{\shortparallel}}{\begin{pmatrix} 2 \\ -1 \end{pmatrix}} \right)$



Set

$$\underline{v}_1 = \underline{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} \underline{v}_2 &= \text{Proj}_{W_1^\perp}(\underline{x}_2) \\ &= \underline{x}_2 - \text{Proj}_{W_1}(\underline{x}_2) \\ &= \underline{x}_2 - \frac{\underline{x}_2 \cdot \underline{v}_1}{\underline{v}_1 \cdot \underline{v}_1} \underline{v}_1 \\ &= \begin{pmatrix} 2 \\ -1 \end{pmatrix} - \frac{\begin{pmatrix} 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ -1 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ -3/2 \end{pmatrix} \end{aligned}$$

Observe

$\text{Span}(\underline{x}_1) = \text{Span}(\underline{v}_1)$ and

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

General Procedure: Build $\{\underline{v}_1, \dots, \underline{v}_p\}$ starting from $\{\underline{x}_1, \dots, \underline{x}_p\}$

Starting from $\underline{v}_1 = \underline{x}_1$ construct $\underline{v}_k$ using $\underline{v}_1, \dots, \underline{v}_{k-1}$ and $\underline{x}_k$ as follows:

1) Let $W_{k-1} = \text{Span}\underbrace{(\underline{v}_1, \dots, \underline{v}_{k-1})}_{\text{orthogonal basis of } W_k}$

2) Let

$$\begin{aligned} \underline{v}_k &= \text{Proj}_{W_{k-1}^\perp}(\underline{x}_k) = \underline{x}_k - \text{Proj}_{W_{k-1}}(\underline{x}_k) \\ &= \underline{x}_k - \frac{\underline{x}_k \cdot \underline{v}_1}{\underline{v}_1 \cdot \underline{v}_1} \underline{v}_1 - \dots - \frac{\underline{x}_k \cdot \underline{v}_{k-1}}{\underline{v}_{k-1} \cdot \underline{v}_{k-1}} \underline{v}_{k-1} \end{aligned}$$

Very Explicitly :

$$\underline{v}_1 = \underline{x}_1$$

$$\underline{v}_2 = \underline{x}_2 - \frac{\underline{x}_2 \cdot \underline{v}_1}{\underline{v}_1 \cdot \underline{v}_1} \underline{v}_1$$

$$\underline{v}_3 = \underline{x}_3 - \frac{\underline{x}_3 \cdot \underline{v}_1}{\underline{v}_1 \cdot \underline{v}_1} \underline{v}_1 - \frac{\underline{x}_3 \cdot \underline{v}_2}{\underline{v}_2 \cdot \underline{v}_2} \underline{v}_2$$

$$\vdots$$

$$\underline{v}_p = \underline{x}_p - \frac{\underline{x}_k \cdot \underline{v}_1}{\underline{v}_1 \cdot \underline{v}_1} \underline{v}_1 - \cdots - \frac{\underline{x}_p \cdot \underline{v}_{p-1}}{\underline{v}_{p-1} \cdot \underline{v}_{p-1}} \underline{v}_{p-1}$$

Facts

1/ $\{\underline{u}_1, \ldots, \underline{u}_p\}$ a basis for W

$\Rightarrow$ $\{\underline{v}_1, \ldots, \underline{v}_p\}$ an orthogonal basis for W

2/ $\text{Span}(\underline{u}_1, \ldots, \underline{u}_k) = \text{Span}(\underline{v}_1, \ldots, \underline{v}_k)$ for all $k \leq p$

3/ To apply Gram-Schmidt Process we must start with a linearly independent set. If $\underline{x}_k$ in span$(\underline{x}_1, \ldots, \underline{x}_{k-1})$

$\Rightarrow$ $\underline{x}_k$ in Span$(\underline{v}_1, \ldots, \underline{v}_{k-1})$ $\Rightarrow$ $\underline{v}_k = 0$ $\Rightarrow$ Cannot proceed further.

Theorem Any subspace $W \subset \mathbb{R}^n$ has an orthogonal basis.

Proof Choose a basis $\{\underline{x}_1, \ldots, \underline{x}_p\}$ for W. Apply the Gram-Schmidt process to get $\{\underline{u}_1, \ldots, \underline{u}_p\}$ an orthogonal basis for W. $\square$

Remark Once we have $\{\underline{v}_1, \ldots, \underline{v}_p\}$ we can alway scale to get an orthonormal basis $\{\underline{u}_1, \ldots, \underline{u}_p\}$ for W.

Interesting Consequence of Gram-Schmidt: QR - Factorization

$$\overset{n\times p \text{ matrix}}{A} = \underbrace{(\underline{x}_1 \cdots \underline{x}_p)}_{\text{L.I.}} \xrightarrow{\text{G.S.}} \overset{n\times p \text{ matrix}}{Q} = \underbrace{(\underline{u}_1 \cdots \underline{u}_p)}_{\text{orthonormal set}}$$

Observe $\text{Span}(\underline{x}_1, \ldots, \underline{x}_k) = \text{Span}(\underline{u}_1, \ldots, \underline{u}_k)$ for $k \le p$

$$\Rightarrow \underline{x}_k = \lambda_{1k}\underline{u}_1 + \lambda_{2k}\underline{u}_2 + \cdots + \lambda_{kk}\underline{u}_k + 0\cdot\underline{u}_{k+1} + \cdots + 0\cdot\underline{u}_p$$

Define

$$R = \begin{array}{c} \begin{pmatrix} \lambda_{11} & \lambda_{12} & & \lambda_{1p} \\ 0 & \lambda_{22} & & \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & & \lambda_{pp} \end{pmatrix} \\ \begin{array}{cccc} \shortparallel & \shortparallel & & \shortparallel \\ \underline{r}_1 & \underline{r}_2 & & \underline{r}_p \end{array} \end{array}$$

$\leftarrow$ $p\times p$ upper triangular matrix

Observe $Q\underline{r}_k = \lambda_{1k}\underline{u}_1 + \lambda_{2k}\underline{u}_2 + \cdots + \lambda_{kk}\underline{u}_k + 0\cdot\underline{u}_{k+1} + \cdots + 0\cdot\underline{u}_p = \underline{x}_k$

$$\Rightarrow QR = (Q\underline{r}_1\ Q\underline{r}_2 \cdots Q\underline{r}_p) = A$$

Remark Columns of Q are orthonormal set $\Rightarrow Q^TQ = I_p$

$$QR = A \Rightarrow Q^TQR = Q^TA \Rightarrow R = Q^TA$$

Conclusion Given A an $n\times p$ matrix with L.I. columns there exists Q an $n\times p$ matrix, with orthonormal columns, and R, a $p\times p$, upper triangular matrix such that

$$A = QR \quad (\Rightarrow R = Q^TA)$$

Example

$$A = \begin{array}{c} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 4 \end{pmatrix} \\ \begin{array}{ccc} \shortparallel & \shortparallel & \shortparallel \\ \underline{x}_1 & \underline{x}_2 & \underline{x}_3 \end{array} \end{array}$$

$$\underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \frac{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\underline{v}_3 = \begin{pmatrix}1\\0\\3\\4\end{pmatrix} - \frac{\begin{pmatrix}1\\0\\3\\4\end{pmatrix}\cdot\begin{pmatrix}1\\0\\0\\0\end{pmatrix}}{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\cdot\begin{pmatrix}1\\0\\0\\0\end{pmatrix}}\begin{pmatrix}1\\0\\0\\0\end{pmatrix} - \frac{\begin{pmatrix}1\\0\\3\\4\end{pmatrix}\cdot\begin{pmatrix}0\\1\\0\\0\end{pmatrix}}{\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\cdot\begin{pmatrix}0\\1\\0\\0\end{pmatrix}}\begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \begin{pmatrix}1\\0\\3\\4\end{pmatrix} - \begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\3\\4\end{pmatrix}$$

$$\underline{u}_1 = \underline{v}_1, \quad \underline{u}_2 = \underline{v}_2, \quad \underline{u}_3 = \begin{pmatrix}0\\0\\3/5\\4/5\end{pmatrix} \qquad \left(\|\underline{v}_3\| = \sqrt{3^2+4^2} = 5 \right)$$

$$\underline{x}_1 = 1\cdot\underline{u}_1 + 0\cdot\underline{u}_2 + 0\cdot\underline{u}_3$$

$$\underline{x}_2 = 1\cdot\underline{u}_1 + 1\cdot\underline{u}_2 + 0\cdot\underline{u}_3$$

$$\underline{x}_3 = 1\cdot\underline{u}_1 + 0\cdot\underline{u}_2 + 5\cdot\underline{u}_3$$

$$\Rightarrow \quad A = \begin{pmatrix}1&1&1\\0&1&0\\0&0&3\\0&0&4\end{pmatrix} = \overset{Q}{\overset{\parallel}{\begin{pmatrix}1&0&0\\0&1&0\\0&0&3/5\\0&0&4/5\end{pmatrix}}}\overset{R}{\overset{\parallel}{\begin{pmatrix}1&1&1\\0&1&0\\0&0&5\end{pmatrix}}}$$