Gram-Schmidt algorithm

Aim lecture: We use the theory of last lecture to give an algorithm for finding orthonormal bases.

As usual in this lecture $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ & V is an $\mathbb{F}$-space equipped with an inner product $(\cdot|\cdot)$.

Gram-Schmidt Algorithm

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ be a basis for V & $V_{\leq r} = \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r)$.

1. The following inductively defines an orthogonal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_d\}$ for V: $\mathbf{w}_1 = \mathbf{v}_1$ &

$$\mathbf{w}_{r+1} = \mathbf{v}_{r+1} - \text{proj}_{V_{\leq r}} \mathbf{v}_{r+1} = \mathbf{v}_{r+1} - \sum_{i=1}^{r} \text{proj}_{\mathbb{F}\,\mathbf{w}_i} \mathbf{v}_{r+1}.$$

2. Furthermore, $\text{Span}(\mathbf{w}_1, \ldots, \mathbf{w}_r) = V_{\leq r}$.

Proof. This is an easy induction. Assuming the statements true for the inner product space $V_{\leq r}$ just note that $\mathbf{w}_{r+1}$ is in $V_{\leq r}^{\perp}$ so spans the orthogonal complement to $V_{\leq r}$ in $V_{\leq r+1}$.

Polynomial example

E.g. Find an orthonormal basis for $V = \mathbb{R}[x]_{\leq 1}$ equipped with the inner product $(f|g) = \int_0^1 \overline{f(t)} g(t) dt$.

Example in $\mathbb{R}^4$

E.g. Find an orthog basis for the image of the following matrix

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

Example cont'd

Projection formula revisited

Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_d\} \subset V$ be an orthogonal set spanning $W \leq V$. Recall for any $\mathbf{v} \in V$ we have

$$\text{proj}_W \mathbf{v} = \frac{(\mathbf{w}_1|\mathbf{v})}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 + \ldots + \frac{(\mathbf{w}_d|\mathbf{v})}{\|\mathbf{w}_d\|^2}\mathbf{w}_d.$$

Prop

Consider the *orthogonal* co-ord system $Q = (\mathbf{w}_1 \ldots \mathbf{w}_d) : \mathbb{F}^d \longrightarrow W$. Let

$$P = \left(\frac{(\mathbf{w}_1|\cdot)}{\|\mathbf{w}_1\|^2}, \ldots, \frac{(\mathbf{w}_d|\cdot)}{\|\mathbf{w}_d\|^2}\right)^T : V \longrightarrow \mathbb{F}^d .$$

1. Then $\text{proj}_W = Q \circ P$.
2. If $V = \mathbb{F}^m$ so $Q \in M_{md}(\mathbb{F})$, then we may re-write $P = D^{-1}Q^*$ where D is the diagonal matrix $D = (\|\mathbf{w}_1\|^2) \oplus \ldots \oplus (\|\mathbf{w}_d\|^2)$. Hence $\text{proj}_W = QD^{-1}Q^*$.

Proof. 1) is just a restatement of the projection formula of lecture 33 above. For 2), just note that the linear map $(\mathbf{w}_i|\cdot)$ is given by left multn by $\mathbf{w}_i^*$ so calculating $D^{-1}C^*$ gives the result.

Example of a projection matrix

E.g. Let $W = \text{Span}((2,1,2)^T, (1,2,-2)^T)$. Find the matrix representing $\text{proj}_W : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$.

Rem Note the formula in prop 2), above generalises the formula $\text{proj}_{\mathbf{u}} = \mathbf{u}\mathbf{u}^T$ for $\mathbf{u} \in \mathbb{R}^n$ a unit vector.

QR-factorisation

Let $A = (\mathbf{v}_1 \dots \mathbf{v}_n) \in M_{mn}(\mathbb{C})$ be an $m \times n$-matrix with lin indep columns $\mathbf{v}_1, \dots, \mathbf{v}_n$. Let $\mathbf{w}_1, \dots, \mathbf{w}_n$ be an orthogonal basis of im A such that $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_i) = \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_i)$ for $i = 1, \dots, n$ e.g. the one that Gram-Schmidt gives, or any derived from that by scaling the vectors.

Prop-Defn [QR-factorisation]

If $Q = (\mathbf{w}_1 \dots \mathbf{w}_n)$ & $D = (\|\mathbf{w}_1\|^2) \oplus \dots \oplus (\|\mathbf{w}_d\|^2)$, then $A = QR$ where $R = D^{-1}Q^*A$ is invertible and upper triangular. When Q has orthonormal columns, we call this a *QR-factorisation* of A.

Proof. Our new projection formula shows $\text{proj}_{\text{im } A} = QD^{-1}Q^*$ so

$$A = (\mathbf{v}_1 \dots \mathbf{v}_n) = (QD^{-1}Q^*\mathbf{v}_1 \dots QD^{-1}Q^*\mathbf{v}_n) = QD^{-1}Q^*A$$

& it suffices to show that $R = D^{-1}Q^*A$ is upper triangular with all diagonal entries non-zero. If $(\beta_1, \dots, \beta_n)^T$ is the i-th column of R then comparing the i-columns of A & QR give

$$\mathbf{v}_i = \beta_1\mathbf{w}_1 + \dots + \beta_n\mathbf{w}_n.$$

Our condition on span ensures $0 = \beta_{i+1} = \beta_{i+2} = \dots = \beta_n$ whilst $\beta_i \neq 0$.

Example

Rem An easy way to remember the formula is to note $Q^*Q = D$ so $Q^*A = Q^*QR = DR$.

E.g. QR-factorise the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$