Gram-Schmidt Process

P. Danziger

1 Orthonormal Vectors and Bases

Definition 1 *A set of vectors* $\{\mathbf{v}_i \mid 1 \le i \le n\}$ *is* orthogonal *if* $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ *whenever* $i \ne j$ *and* orthonormal *if* $\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$

For ease of notation, we define the the Kronecker delta function δ_{ij} to be the discrete function $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$. The matrix associated with the Kronecker delta, $[\delta_{ij}] = I$.
Note that an orthonormal set is an orthogonal set of unit vectors.

Definition 2 *An* Orthonormal Basis *is an orthonormal set of vectors, which is also a basis.*

Example 3

1. The standard basis in $\mathbb{R}^2$. $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1) \in \mathbb{R}^2$.

$$\begin{aligned} \mathbf{e}_1 \cdot \mathbf{e}_1 &= (1,0)\cdot(1,0) = 1, \\ \mathbf{e}_2 \cdot \mathbf{e}_2 &= (0,1)\cdot(0,1) = 1, \\ \mathbf{e}_1 \cdot \mathbf{e}_2 &= \mathbf{e}_2 \cdot \mathbf{e}_1 = (1,0)\cdot(0,1) = 0. \end{aligned}$$

   So $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 1 & i = j \quad (i = j = 1 \text{ or } i = j = 2) \\ 0 & i \ne j \quad (i = 1, j = 2 \text{ or } i = 2, j = 1) \end{cases}$

2. The standard basis in $\mathbb{R}^n$. $\{\mathbf{e}_i \in \mathbb{R}^n \mid 1 \le i \le n\}$.

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$$

3. $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1,-1) \in \mathbb{R}^2$.

$$\begin{aligned} \mathbf{u}_1 \cdot \mathbf{u}_1 &= \tfrac{1}{\sqrt{2}}(1,1)\cdot\tfrac{1}{\sqrt{2}}(1,1) = \tfrac{1}{2}(1+1) = 1 \\ \mathbf{u}_2 \cdot \mathbf{u}_2 &= \tfrac{1}{\sqrt{2}}(1,-1)\cdot\tfrac{1}{\sqrt{2}}(1,-1) = \tfrac{1}{2}(1+(-1)^2) = 1, \\ \mathbf{u}_1 \cdot \mathbf{u}_2 &= (1,1)\cdot(1,-1) = \tfrac{1}{2}(1-1) = 0. \end{aligned}$$

   So $\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij}$.

4. $\mathbf{u}_1 = (0,1,0)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1,0,-1)$, $\mathbf{u}_3 = \frac{1}{\sqrt{2}}(1,0,1) \in \mathbb{R}^3$.

$$\begin{aligned}
\mathbf{u}_1 \cdot \mathbf{u}_1 &= (0,1,0)\cdot(0,1,0) = 1\\
\mathbf{u}_2 \cdot \mathbf{u}_2 &= \tfrac{1}{\sqrt{2}}(1,0,-1)\cdot\tfrac{1}{\sqrt{2}}(1,0,-1) = \tfrac{1}{2}(1+(-1)^2) = 1,\\
\mathbf{u}_3 \cdot \mathbf{u}_2 &= \tfrac{1}{\sqrt{2}}(1,0,1)\cdot\tfrac{1}{\sqrt{2}}(1,0,1) = \tfrac{1}{2}(1+1) = 1,\\
\mathbf{u}_1 \cdot \mathbf{u}_2 &= (0,1,0)\cdot\tfrac{1}{\sqrt{2}}(1,0,-1) = 0.\\
\mathbf{u}_1 \cdot \mathbf{u}_3 &= (0,1,0)\cdot\tfrac{1}{\sqrt{2}}(1,0,1) = 0.\\
\mathbf{u}_2 \cdot \mathbf{u}_3 &= \tfrac{1}{\sqrt{2}}(1,0,-1)\cdot\tfrac{1}{\sqrt{2}}(1,0,1) = \tfrac{1}{\sqrt{2}}(1=(-1)) = 0.
\end{aligned}$$

So $\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij}$.

Theorem 4 *Given an orthonormal set $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ of vectors in $\mathbb{R}^n$ and an arbitrary vector $\mathbf{v} \in \text{span}(B)$ the components of $\mathbf{u}$ with respect to V are given by*

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\,\mathbf{u_1} + (\mathbf{v}\cdot\mathbf{u}_2)\,\mathbf{u_2} + \ldots + (\mathbf{v}\cdot\mathbf{u}_k)\,\mathbf{u_k}$$

Proof: Proof: Since the $\mathbf{u}_i$ are unit, $||\mathbf{u}_j|| = 1$, and so $\text{proj}_{\mathbf{u}_i}\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{u}_i}{||\mathbf{u}_j||^2}\mathbf{u}_j = \mathbf{v}\cdot\mathbf{u}_i$ gives the projection of $\mathbf{u}$ in the direction $\mathbf{v}_i$. Since $\mathbf{v}_i$ is orthogonal to $\mathbf{v}_j$ $(i \neq j)$, $\text{proj}_{\mathbf{v}_i}\mathbf{u}$ is orthogonal to $\text{proj}_{\mathbf{v}_j}\mathbf{u}$ $\square$
If $k = n$, so V is an orthonormal basis for $\mathbb{R}^n$, then we can write the components of $\mathbf{v}$ with respect to the basis V as

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1,\ \mathbf{v}\cdot\mathbf{u}_2,\ \ldots,\ \mathbf{v}\cdot\mathbf{u}_n)_V$$

Example 5 *Find the components of $\mathbf{v} = (2,1,1)$ with respect to the orthonormal basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, where*

$$\mathbf{u}_1 = (0,1,0),\ \mathbf{u}_2 = \frac{1}{\sqrt{2}}(1,0,-1),\ \mathbf{u}_3 = \frac{1}{\sqrt{2}}(1,0,1) \text{ of } \mathbb{R}^3.$$

$$\begin{aligned}
\mathbf{v}\cdot\mathbf{u}_1 &= (2,1,1)\cdot(0,1,0) = 1\\
\mathbf{v}\cdot\mathbf{u}_2 &= (2,1,1)\cdot\tfrac{1}{\sqrt{2}}(1,0,-1) = \tfrac{1}{\sqrt{2}}(2-1) = \tfrac{1}{\sqrt{2}}\\
\mathbf{v}\cdot\mathbf{u}_3 &= (2,1,1)\cdot\tfrac{1}{\sqrt{2}}(1,0,1) = \tfrac{1}{\sqrt{2}}(2+1) = \tfrac{3}{\sqrt{2}}
\end{aligned}$$

So $\mathbf{v} = \mathbf{u}_1 + \frac{1}{\sqrt{2}}\mathbf{u}_2 + \frac{1}{\sqrt{2}}\mathbf{u}_3$. Or $\mathbf{v} = (1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})_B$.

Theorem 6 *A basis B is an orthonormal basis if and only if the corresponding change of basis matrix A_B is orthogonal.*

Proof:

($\Rightarrow$) Let $B = \{\mathbf{v}_i \mid 1 \leq i \leq n\}$ be an orthonormal basis and A_B be its corresponding change of basis matrix. The columns of A_B are the vectors of B, so, noting that the rows of A_B^T are the vectors of B,

$$A_B^T A_B = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix} = [\mathbf{v}_i\cdot\mathbf{v}_j] = [\delta_{ij}] = I$$

So $(A_B)^T A_B = I$ and so A_B is orthogonal.

($\Leftarrow$) Suppose A is an orthogonal matrix, so $A^T = A^{-1}$, and let $B = \{\mathbf{v}_i \mid 1 \le i \le n\}$ be the columns of A. We wish to show that B is an orthonormal basis. Since A is invertible, $\det(A) \neq 0$ and so the columns of A are linearly independent and form a basis. Now, as above,

$$[\mathbf{v}_i \cdot \mathbf{v}_j] = A_B^T A_B = I = [\delta_{ij}].$$

So B is an orthonormal basis

2 The Gram-Schmidt Procedure

Given an arbitrary basis we can form an orthonormal basis from it by using the 'Gram-Schmidt Process'. The idea is to go through the vectors one by one and subtract off that part of each vector that is not orthogonal to the previous ones. Finally, we make each vector in the resulting basis unit by dividing it by its norm.

Gram-Schmidt Procedure
Given a basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ of $\mathbb{R}^n$, produce a new basis $B' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of $\mathbb{R}^n$ which is orthonormal. In the process we produce a third basis $B'' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ of $\mathbb{R}^n$ whose members are orthogonal to each other, but not unit.

1. Set $\mathbf{w}_1 = \mathbf{u}_1$.
2. For each i from 2 to n

   (a) Set $$\begin{aligned}\mathbf{w}_i &= \mathbf{u}_i - \textstyle\sum_1^{i-1} \mathrm{proj}_{\mathbf{w}_j} \mathbf{u}_i \\ &= \mathbf{u}_i - \textstyle\sum_1^{i-1} \dfrac{\mathbf{u}_i \cdot \mathbf{w}_j}{||\mathbf{w}_j||^2} \mathbf{w}_j\end{aligned}$$

   We now have an orthogonal basis $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ of $\mathbb{R}^n$. The final step makes all members of the basis unit.
3. For each i from 1 to n set $\mathbf{v}_i = \frac{1}{||\mathbf{w}_i||}\mathbf{w}_i$.

So when $n = 3$

1. Set $\mathbf{w}_1 = \mathbf{u}_1$.
2. (a) $i = 2$
   $$\mathbf{w}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2}.$$
   (b) $i = 3$
   $$\mathbf{w}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_2 \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2} - \frac{\mathbf{u}_2 \cdot \mathbf{w}_2}{||\mathbf{w}_2||^2}.$$
3. $$\mathbf{v}_1 = \frac{1}{||\mathbf{w}_1||}\mathbf{w}_1, \quad \mathbf{v}_2 = \frac{1}{||\mathbf{w}_2||}\mathbf{w}_2, \quad \mathbf{v}_3 = \frac{1}{||\mathbf{w}_3||}\mathbf{w}_3.$$

Example 7

1. Given $B = \{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{u}_1 = (0, 1)$ and $\mathbf{u}_2 = (1, 1)$, use the Gram-Schmidt procedure to find a corresponding orthonormal basis.

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{u}_1 = (0, 1), \\
\mathbf{w}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2}\mathbf{w}_1 \\
&= (1, 1) - \frac{(1, 1) \cdot (0, 1)}{||(0, 1)||}(0, 1) \\
&= (1, 1) - (0, 1) \\
\mathbf{w}_2 &= (1, 0)
\end{aligned}$$

Now $\mathbf{w}_1$ and $\mathbf{w}_2$ are already unit, so $\mathbf{v}_1 = \mathbf{w}_1$ and $\mathbf{v}_2 = \mathbf{w}_2$.

2. Given $B = \{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{u}_1 = (1, 1)$ and $\mathbf{u}_2 = (2, 1)$, use the Gram-Schmidt procedure to find a corresponding orthonormal basis.

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{u}_1 = (1, 1), \\
\mathbf{w}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2}\mathbf{w}_1 \\
&= (2, 1) - \frac{(2, 1) \cdot (1, 1)}{||(1, 1)||^2}(1, 1) \\
&= (2, 1) - \tfrac{3}{2}(1, 1) \\
&= (\tfrac{1}{2}, -\tfrac{1}{2}) \\
\mathbf{w}_2 &= \tfrac{1}{2}(1, -1)
\end{aligned}$$

Now $\mathbf{v}_1 = \frac{1}{||\mathbf{w}_1||}\mathbf{w}_1 = \frac{1}{\sqrt{2}}(1, 1)$ and $\mathbf{v}_2 = \frac{1}{||\mathbf{w}_2||}\mathbf{w}_2 = \frac{1}{\sqrt{2}}(1, -1)$ is an orthonormal basis.

3. Given $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, where $\mathbf{u}_1 = (1, 2, 1)$, $\mathbf{u}_2 = (1, 1, 3)$ and $\mathbf{u}_3 = (2, 1, 1)$, use the Gram-Schmidt procedure to find a corresponding orthonormal basis.

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{u}_1 = (1, 2, 1), \\
\mathbf{w}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2}\mathbf{w}_1 \\
&= (1, 1, 3) - \frac{(1, 1, 3) \cdot (1, 2, 1)}{||(1, 2, 1)||^2}(1, 2, 1) \\
&= (1, 1, 3) - \tfrac{6}{6}(1, 2, 1) \\
\mathbf{w}_2 &= (0, -1, 2) \\
\mathbf{w}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2}\mathbf{w}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{w}_2}{||\mathbf{w}_2||^2}\mathbf{w}_2 \\
&= (2, 1, 1) - \frac{(2, 1, 1) \cdot (1, 2, 1)}{||(1, 2, 1)||^2}(1, 2, 1) - \frac{(2, 1, 1) \cdot (0, -1, 2)}{||(0, -1, 2)||^2}(0, -1, 2) \\
&= (2, 1, 1) - \tfrac{5}{6}(1, 2, 1) - \tfrac{1}{5}(0, -1, 2) \\
\mathbf{w}_2 &= \tfrac{1}{30}(35, -14, 17)
\end{aligned}$$

Now

$$\begin{aligned}
\mathbf{v}_1 &= \tfrac{1}{||\mathbf{w}_1||}\mathbf{w}_1 = \tfrac{1}{\sqrt{6}}(1,2,1), \\
\mathbf{v}_2 &= \tfrac{1}{||\mathbf{w}_2||}\mathbf{w}_2 = \tfrac{1}{\sqrt{5}}(0,-1,2) \\
\mathbf{v}_3 &= \tfrac{1}{||\mathbf{w}_3||}\mathbf{w}_3 = \tfrac{1}{\sqrt{57}}(35,-14,17)
\end{aligned}$$

is an orthonormal basis.

Check $\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}$.