5.3 Gram-Schmidt

By now we are used to finding bases, but recall that orthogonal, or even better, orthonormal bases are preferred.

Definition 1. The vectors $q_1, \ldots, q_n$ are orthonormal if

$$q_i^T q_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

(giving orthogonality),

$$q_i^T q_i = 1$$

(giving the normalization);

We can also create matrices out of these bases. Notice that the standard basis for an Euclidean space is in the columns of the identity matrix. However, if we want a generic orthonormal basis we need to apply the Gram-Schmidt orthogonalization procedure.

Theorem 1. If $Q$ (square or rectangular) has orthonormal columns, then $Q^T Q = I$.

Definition 2. An orthogonal matrix is a square matrix with orthonormal columns.

Theorem 2. For orthogonal matrices, the transpose is the inverse.

Ex: Consider

$$Q = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

which we can verify by multiplying.

Ex: Any permutation matrix $P$ (consisting of only row exchanges) is an orthogonal matrix. The criteria of orthonormal columns and square are trivially satisfied. Then we check $P^{-1} = P^T$ by checking $P P^T = I$.

Theorem 3. Multiplication by any $Q$ preserves lengths: $||Qx|| = ||x||$ for all $x$.

It also preserves inner products and angles: $(Qx)^T (Qy) = x^T Q^T Q y = x^T y$.

Consider $Qx = b$ where $q_i$ are the columns of $Q$. Then we can write

$$b = x_1 q_1 + x_2 q_2 + \cdots + x_n q_n$$

If we multiply both sides by $q_i^T$ we get

$$q_i^T b = q_i^T (x_1 q_1 + x_2 q_2 + \cdots + x_n q_n) = q_i^T x = x_i \Rightarrow x = Q^T b.$$ So if your $A$ is an orthogonal matrix, you don’t have to do Gaussian Elimination.

The Gram-Schmidt Process

Suppose you are given three independent vectors $\vec{a}, \vec{b}, \vec{c}$. If they are orthonormal we can project a vector $\vec{v}$ onto $\vec{a}$ by doing $(\vec{a}^T \vec{v}) \vec{a}$. To project onto the $\vec{a} - \vec{b}$ plane we do $(\vec{a}^T \vec{v}) \vec{a} + (\vec{b}^T \vec{v}) \vec{b}$, etc.

Process: We are given $\vec{a}, \vec{b}, \vec{c}$ and we want $q_1, q_2, q_3$. No problem with $q_1$; i.e., $q_1 = a/||a||$ (we don’t have to change its direction, just normalize.) The problem begins with $q_2$, which has to be orthogonal to $q_1$. If the vector $b$ has any component in the direction of $q_1$ (i.e., direction of $a$) it has to be subtracted: $B = b - (q_1^T b) q_1$, then $q_2 = B/||B||$, and this continues for $q_3$: $C = c - (q_1^T c) q_1 - (q_2^T c) q_2$, then $q_3 = C/||C||$, so on and so forth.

Ex: $a = (1,0,1), b = (1,0,0), and c = (2,1,0)$ for $A = [a \ b \ c]$.

Solution:

Step 1: Make the first vector into a unit vector: $q_1 = a/\sqrt{2} = (1/\sqrt{2}, 0, 1/\sqrt{2})$.

Step 2a: Subtract from the second vector its component in the direction of the first: $B = b - (q_1^T b) q_1 = (1/2, 0, -1/2)$.

Step 2b: Divide $B$ by its magnitude: $q_2 = B/||B|| = (1/\sqrt{2}, 0, -1/\sqrt{2})$.

Step 3a: Subtract from the third vector its component in the first and second directions: $C = c - (q_1^T c) q_1 - (q_2^T c) q_2 = (0, 1, 0)$.

Step 3b: We normalize $C$, but $C$ is already a unit vector so $q_3 = (0, 1, 0)$.

Then we can write $Q$ as the matrix

$$Q = \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}.$$

From the matrix $Q$ we can get a $A = Q R$ factorization. This means that $A = Q R$.

$$A = \begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} q_1^T a & q_1^T b & q_1^T c \\ q_2^T a & q_2^T b & q_2^T c \\ q_3^T a & q_3^T b & q_3^T c \end{pmatrix} \begin{pmatrix} q_1^T b & q_1^T c \\ q_2^T b & q_2^T c \\ q_3^T b & q_3^T c \end{pmatrix}.$$
Now let’s do a bunch of examples from the book on page 263.

1) They are orthogonal but not normal.

5) Same as above.

25) We can get $q_1 = (3, 4)/5$ immediately. Then
   \[
   B = b - (q_1^T b)q_1 = (1, 0) - \frac{3}{5}(3, 4)/5 = (16/25, -12/25) \Rightarrow q_2 = \frac{(4^2/5^2, -12/5^2)}{\sqrt{(4^2/5^2) + (3^2 \cdot 4^2)/5^2}} = (4/5, -3/5)
   \]

27) $q_1 = (0, 1)$. Then
   \[
   B = b - (q_1^T b)q_1 = (2, 5) - 5(0, 1) = (2, 0) \Rightarrow q_2 = (1, 0).
   \]

29) The vectors are already orthogonal, so just divide by the magnitude.

33) $q_1 = (0, 1, 1)/\sqrt{2}$. Then
   \[
   B = b - (q_1^T b)q_1 = (1, 1, 0) - \frac{1}{\sqrt{2}}(0, 1/\sqrt{2}, 1/\sqrt{2}) = (1, 1/2, -1/2) \Rightarrow q_2 = (1, 1/2, -1/2)/\sqrt{3/2} = (\sqrt{2/3}, \sqrt{2/3}/2, -\sqrt{2/3}/2).
   \]

And
   \[
   C = c - (q_1^T c)q_1 - (q_2^T c)q_2 = (1, 0, 1) - \frac{1}{\sqrt{2}}(0, 1/\sqrt{2}, 1/\sqrt{2}) - \frac{\sqrt{2/3}}{2}(\sqrt{2/3}, \sqrt{2/3}/2, -\sqrt{2/3}/2)
   \]
   \[
   = (1, 0, 1) - (0, 1/2, 1/2) - (1/3, 1/6, -1/6) = (2/3, -2/3, 2/3).
   \]

So,
   \[
   q_3 = (2/3, -2/3, 2/3)/(2\sqrt{2/3}) = (1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}).
   \]