Professor Jennifer Balakrishnan, *jbala@bu.edu*

What is on today

1 The Gram-Schmidt process

Lay–Lay–McDonald §6.4 pp. 356 – 360

The Gram-Schmidt algorithm produces an orthogonal basis for any nonzero subspace of $\mathbb{R}^n$. We illustrate it with a few examples and then state the algorithm as a theorem.

Example 1. *Let* $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$*, where* $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ *and* $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. *Construct an orthogonal basis* $\{\mathbf{v}_1, \mathbf{v}_2\}$ *for* W.

Example 2. *Let* $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. *Then* $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ *is linearly independent and is a basis for a subspace* W *of* $\mathbb{R}^4$. *Construct an orthogonal basis for* W.

Theorem 3 (Gram-Schmidt Orthogonalization). *Given a basis* $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ *for a nonzero subspace* W *of* $\mathbb{R}^n$*, define*

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \cdots \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}.
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ *is an orthogonal basis for* W. *In addition* $\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ *for* $1 \le k \le p$.

An orthonormal basis is easily constructed from an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$: just normalize (rescale) all of the $\mathbf{v}_k$. When doing these computations, it's easier to normalize the full basis at the end rather than each individual vector as soon as it is found.

If an $m \times n$ matrix A has linearly independent columns $\mathbf{x}_1, \ldots, \mathbf{x}_n$, then applying Gram-Schmidt (with normalizations) to $\mathbf{x}_1, \ldots, \mathbf{x}_n$ amounts to factoring A, as described in the next theorem. This factorization is widely used in algorithms for solving equations and finding eigenvalues.

Theorem 4 (QR factorization). *If* A *is an* $m \times n$ *matrix with linearly independent columns, then* A *can be factored as* $A = QR$*, where* Q *is an* $m \times n$ *matrix whose columns form an orthonormal basis for* $\mathrm{Col}\, A$ *and* R *is an* $n \times n$ *upper triangular invertible matrix with positive entries on its diagonal.*

Example 5. *Find a QR factorization of* $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.