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 $\mathbb{R}^n$.

Theorem 1 (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.

Example 2. *Let* $W = \mathrm{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$, *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.

$v_1 = x_1 = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$

$v_2 = x_2 - \frac{x_2 \cdot v_1}{\underline{\underline{v_1 \cdot v_1}}} v_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}}{\left\|\begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}\right\|^2} \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{15}{45}\begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$

$= \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$

$= \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$

So $\left\{ \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \right\}$ is an orthogonal basis for W.

Example 3. *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.

$$\boxed{v_1 = x_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$

rescale to clear denominators $(\cdot 4)$ $\rightsquigarrow$ $\boxed{\begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix} =: v_2'}$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2'}{v_2' \cdot v_2'} v_2' = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{12}\begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

(note, just using rescaled v_2)

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1/2 \\ -1/2 \\ -1/2 \\ -1/2 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ -1/6 \\ -1/6 \\ -1/6 \end{pmatrix} = \begin{pmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{pmatrix} \overset{\text{rescale}}{\rightsquigarrow} \boxed{\begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}}$$

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* Col A *and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.*

* Q comes from Gram-Schmidt + normalization

Then $Q^T Q = I$

Since $A = QR$

* $Q^T A = Q^T Q R = I R = R$.

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

We saw in previous example that

$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$ gave orthogonal basis.

Now compute norms:

$\|v_1\| = 2$, $\|v_2\| = \sqrt{12}$, $\|v_3\| = \sqrt{6}$

$$Q = \left[\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \frac{v_3}{\|v_3\|} \right]$$

$$\boxed{Q = \begin{pmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{pmatrix}}$$

Use $Q^T A = R$ to get

$$\underbrace{\begin{pmatrix} 1/2 & 1/2 & \frac{1}{2} & \frac{1}{2} \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & \frac{1}{\sqrt{6}} \end{pmatrix}}_{3 \times 4} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{4 \times 3}$$

$$\boxed{R = \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{pmatrix}}$$