What is on today

1 The Gram-Schmidt process

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 $\{x_1, \ldots, x_p\}$ for a nonzero subspace $W$ of $\mathbb{R}^n$, define

\[
\begin{align*}
    v_1 &= x_1 \\
v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
    &\vdots \\
v_p &= x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}.
\end{align*}
\]

Then $\{v_1, \ldots, v_p\}$ is an orthogonal basis for $W$. In addition $\text{Span}\{v_1, \ldots, v_k\} = \text{Span}\{x_1, \ldots, x_k\}$ for $1 \leq k \leq p$.

An orthonormal basis is easily constructed from an orthogonal basis $\{v_1, \ldots, v_p\}$: just normalize (rescale) all of the $v_k$. 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 $W = \text{Span}\{x_1, x_2\}$, where $x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construct an orthogonal basis $\{v_1, v_2\}$ for $W$. 

Example 3. Let \( x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \). Then \( \{x_1, x_2, x_3\} \) is linearly independent and is a basis for a subspace \( W \) of \( \mathbb{R}^4 \). Construct an orthogonal basis for \( W \).

If an \( m \times n \) matrix \( A \) has linearly independent columns \( x_1, \ldots, x_n \), then applying Gram-Schmidt (with normalizations) to \( x_1, \ldots, x_n \) 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 \( \text{Col} \ A \) and \( R \) is an \( n \times n \) upper triangular invertible matrix with positive entries on its diagonal.

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