QR Decomposition

- The QR decomposition is perhaps the most important algorithmic idea in numerical linear algebra.
  - ◇ Suppose $A \in R^{m\times n}$, $m \geq n$, and suppose rank $(A) = n$ (i.e., suppose A has linearly independent columns).
  - ◇ The matrix A can always be decomposed as the product

$$A = QR, \tag{0.1}$$

    where $Q = R^{m\times m}$ is an orthogonal matrix (i.e., $Q^TQ = I_m$), and $R \in R^{m\times n}$ is a matrix of the form $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ with $R_1 \in R^{n\times n}$ an upper triangular matrix.
    - ▷ Because of the zeros at the lower part of R, the decomposition can be simplified as

$$A = Q_1R_1 \tag{0.2}$$

      where $Q_1 \in R^{m\times n}$ is made of the first n columns of Q and satisfies $Q_1^TQ_1 = I_n$ (but it is not orthogonal.) This is called the *reduced QR factorization* of A.
- Suppose $A = QR$. Then the system $Ax = b$ can be solved from the triangular system $Rx = Q^Tb$. This is an excellent method for solving linear systems, but it is *not* the standard method for such problems.
  - ◇ Gaussian elimination requires only half as many numerical operations in the QR decomposition.
- The QR decomposition can also be used in solving least squares problems and eigenvalue problems.

Gram-Schmidt Orthogonalization

- Given a set of linearly independent column vectors $\{v_1, \ldots, v_n\} \subset R^m$ $(m \geq n)$, these vectors span an n-dimensional vector subspace of R^m. We could construct an orthonormal set of vectors $\{q_1, \ldots, q_n\}$ that spans the same subspace. The well-known *Gram-Schmidt process* provides a recipe for accomplishing this:
  - ◇ Set $w_1 := v_1$.
  - ◇ Set, in general, for $j = 2, \ldots, n$

$$w_j := v_j - \sum_{i=1}^{j-1} \frac{\langle w_i, v_j \rangle}{\langle w_i, w_i \rangle} w_i.$$

    - ▷ It is really to be verified that $\langle w_i, w_j \rangle = 0$ for $i \neq j$, and that $w_j \neq 0$.
    - ▷ If we define $q_j := \frac{w_j}{\|w_j\|_2}$, then $\langle q_i, q_j \rangle = \delta_{ij}$.

- The above relationship may be rewritten as

$$\begin{aligned} v_1 &= w_1; \\ v_j &= w_j + \sum_{i=1}^{j-1} \frac{\langle w_i, v_j \rangle}{\langle w_i, w_i \rangle} w_i. \end{aligned}$$

- In matrix form, we may record the above relationship as

$$\begin{aligned} & [v_1, \ldots, v_n] \\ = \; & [w_1, \ldots, w_j, \ldots, w_n] \begin{bmatrix} 1 & \frac{\langle w_1, v_2 \rangle}{\langle w_i, w_1 \rangle} & \frac{\langle w_1, v_3 \rangle}{\langle w_1, w_1 \rangle} & & \\ 0 & 1 & \frac{\langle w_2, v_3 .}{\langle w_2, w_2 \rangle} & & \\ 0 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}. \end{aligned}$$

  That is, $V = W\tilde{R}$.

- ⋄ Let $D :=$ diag $\{\frac{1}{\|w_1\|}, \ldots, \frac{1}{\|w_n\|}\}$ denote the scaling factors.
- ⋄ Then $V = WDD^{-1}\tilde{R} = Q_1R_1$ where $Q_1 := WD$ and $R_1 := D^{-1}\tilde{R}$.

- For our consideration, identify the matrix V as our original matrix A.
  - ⋄ The Gram-Schmidt process offers a numerical procedure to compute the QR decomposition.
  - ⋄ In practice, the sequence of Gram-Schmidt calculations turns to be numerically unstable. Fortunately, there are better way.
  - ⋄ The insight we obtained from the Gram-Schmidt process is that the columns of A and columns of Q generate the same range space.