Two QR Factorizations

We compare two techniques for QR factorizations of a full-rank matrix $A \in \mathbb{R}^{m \times n}$, with $m \geq n$. While there are a few other methods available for use, we will talk here about the *modified Gram-Schmidt* process (MGS), and the Householder QR factorization (HQR).

MGS *thin* QR factorization:

$$A = QR, \qquad \text{where}$$

$Q \in \mathbb{R}^{m \times n}$ satisfies $Q^T Q = I$ and $R \in \mathbb{R}^{n \times n}$ is upper triangular. The cost is $2mn^2 + \mathrm{O}(mn)$ flops. If A is overwritten by Q, then only $\frac{1}{2}n^2 + \mathrm{O}(n)$ words of memory are required. If $\tilde{Q}$ and $\tilde{R}$ are the computed versions of Q and R, then there exists $\delta A \in \mathbb{R}^{m \times n}$ with $A + \delta A = \tilde{Q}\tilde{R}$, where $\|\delta A\| = \|A\|\mathrm{O}(\mu)$, and $\|\tilde{Q}^T \tilde{Q} - I\| = \kappa(A)\mathrm{O}(\mu)$.

HQR factored-Q *full* QR factorization:

$$A = QR, \qquad \text{where}$$

$Q \in \mathbb{R}^{m \times m}$ satisfies $Q^T Q = QQ^T = I$ and $R \in \mathbb{R}^{m \times n}$ is upper triangular. We say "factored" here because HQR does not produce Q, but instead produces $u_1, u_2, \ldots, u_n$, where $H_k = H(u_k)$ and $Q = H_1 H_2 \cdots H_s$. The cost is $2mn^2 - \frac{2}{3}n^3 + \mathrm{O}(mn)$ flops. If A is overwritten by $u_1, u_2, \ldots, u_n$ and the strict upper triangle of R (for example), then only $\mathrm{O}(n)$ words of memory are required. If $\tilde{R}$ is the computed version of R and $\tilde{Q}$ is the *exactly formed* Q matrix defined by the *computed* $u_1, u_2, \ldots, u_n$, then there exists $\delta A \in \mathbb{R}^{m \times n}$ with $A + \delta A = \tilde{Q}\tilde{R}$, where $\|\delta A\| = \|A\|\mathrm{O}(\mu)$.

HQR explicit-Q *full* QR factorization:

Let's say $Q = [Q_1 \;\; Q_2]$, where $Q_1 \in \mathbb{R}^{m \times n}$. If only Q_1, is needed, then the flop requirements are doubled, to $4mn^2 - \frac{4}{3}n^3$, and the memory requirements are $\frac{1}{2}mn + \mathrm{O}(n)$. If $\bar{Q}_1$ is the computed version of Q_1, then $\|\bar{Q}_1^T \bar{Q}_1 - I\| = \mathrm{O}(\mu)$. If all of Q is explicitly required, then the flop requirements become $4m^2 n - 2mn^2 + \frac{2}{3}n^3$ and memory requirements become $\mathrm{O}(m^2)$ words. If $\bar{Q}$ is that computed version of $\tilde{Q}$, then $\|\bar{Q}^T \bar{Q} - I\| = \mathrm{O}(\mu)$.

MGS & HQR

both represent a orthonormal basis (O.B.) for $\mathrm{ColSp}(A)$. In exact arithmetic, each column of Q from MGS is $\pm$ the corresponding column of Q_1 from HQR. In other words, the thin part of the full QR is the thin QR. MGS computes Q_1 faster, but explicit HQR gives a "more orthogonal" basis. Implemented with care, both methods are backward stable for (LS).