Differentiating the qr decomposition

James Townsend

December 17, 2018

The QR decomposition of an $m \times n$ matrix A (with $m \geq n$) is defined by

$$A = QR \tag{1}$$

where Q is an $m \times n$ matrix with orthonormal columns, and R is an $n \times n$ upper triangular matrix.

The orthonormality of the columns of Q implies that

$$Q^\top Q = I. \tag{2}$$

Differentiating this leads to

$$dQ^\top Q + Q^\top dQ = 0 \tag{3}$$

which implies that the matrix $dQ^\top Q$ is skew-symmetric. By counting parameters, it can be shown that dQ can be uniquely decomposed into two terms:

$$dQ = Qd\Omega + Q_\perp dK \tag{4}$$

where $Q_\perp$ is an $m \times (m-n)$ matrix such that $\begin{bmatrix} Q & Q_\perp \end{bmatrix}$ is an orthogonal matrix (this could be computed using the Gram-Schmidt process), $d\Omega$ is $n \times n$ skew-symmetric and dK is an unconstrained $(m-n) \times n$ matrix.

The derivative of R is also constrained. Since R is upper triangular dR must also be upper triangular.

Differentiating equation 1 gives

$$dA = dQR + QdR. \tag{5}$$

Substituting the expression for dQ from (4) into this equation gives

$$\begin{aligned} dA &= [Qd\Omega + Q_\perp dK]\, R + QdR & (6) \\ &= Q\, [d\Omega R + dR] + Q_\perp dKR. & (7) \end{aligned}$$

Left multiply this expression by $Q^\top$, and right multiply by R^{-1}. This gives

$$Q^\top dAR^{-1} = d\Omega + dRR^{-1} \tag{8}$$

The right hand side of this expression is the sum of a skew-symmetric matrix ($d\Omega$) and an upper triangular one (dRR^{-1}), thus each term is uniquely deterimined by the left hand side.

To find dK, multiply (7) by ${Q_\perp}^\top$ on the left and R^{-1} on the right leading to

$$dK = {Q_\perp}^\top dAR^{-1}. \tag{9}$$