Derivative of the Polar Decomposition

Julian Panetta

July 5, 2018

We consider the derivative of the polar decomposition of a "time-varying" matrix $A(t)$:

$$A(t) = R(t)S(t) \tag{1}$$

where R is the closest rotation matrix to A (in Frobenius norm) and S is a symmetric matrix capturing the stretching performed by A. Notice that if A inverts space (if $\det(A) < 0$), then this decomposition is actually *not* the polar decomposition but rather the closely related decomposition where the smallest eigenvalue of S is negated.

We begin by differentiating both sides of (1), denoting quantities' instantaneous rates of change with a dot (e.g., $\dot{A} := \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} A(t)$):

$$\dot{A} = \dot{R}S + R\dot{S} \quad \Longleftrightarrow \quad R^T\dot{A} = R^T\dot{R}S + \dot{S}. \tag{2}$$

We observe that because $S(t)$ is symmetric for all t, $\dot{S}$ must be a symmetric matrix. Furthermore, because $R^TR = I$, we find by differentiating both sides that $R^T\dot{R}$ is a skew symmetric matrix: $\dot{R}^TR + R^T\dot{R} = 0$. We can use these symmetric and skew symmetric properties to isolate and solve for $\dot{R}$; computing the skew symmetric part of both sides kills off the symmetric $\dot{S}$ term:

$$\begin{aligned} R^T\dot{A} - \left(R^T\dot{A}\right)^T &= R^T\dot{R}S - \left(R^T\dot{R}S\right)^T + \cancel{\dot{S} - \dot{S}^T} \\ \underbrace{R^T\dot{A} - \dot{A}^TR}_{C} &= \underbrace{R^T\dot{R}}_{M}S - S\left(R^T\dot{R}\right)^T = MS - SM^T, \end{aligned}$$

which is a *Sylvester equation* $C = MS + SM$ for skew symmetric matrix $M := R^T\dot{R}$. We solve this equation by using S's eigen decomposition $S = Q\Lambda Q^T$:

$$C = MQ\Lambda Q^T + Q\Lambda Q^TM \quad \Longleftrightarrow \quad Q^TCQ = Q^TMQ\Lambda + \Lambda Q^TMQ.$$

It becomes clear that this equation is now easy to solve when we inspect its components:

$$[Q^TCQ]_{ij} = [Q^TMQ]_{ij}\lambda_j + \lambda_i[Q^TMQ]_{ij} = (\lambda_i + \lambda_j)[Q^TMQ]_{ij}. \tag{3}$$

(In this equation i, j are free indices and summation is *not* implied.) In other words, we simply divide the ij^{th} component of Q^TCQ by $\lambda_i + \lambda_j$ to find Q^TMQ. From here we are essentially done: we can just compute $\dot{R} = RM = RQ(Q^TMQ)Q^T$.

We can express the division operation from (3) in terms of standard linear algebra operations by converting matrix equation (3) into a vector equation. First, we notice that both Q^TCQ and Q^TMQ are skew symmetric matrices because both C and M are. Any skew symmetric matrix B has the form

$$B = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix},$$

where we chose the signs so that B is actually the cross product matrix for vector $\mathbf{b} = (b_1, b_2, b_3)^T$. In other words, we have the property that $B\mathbf{v} = \mathbf{b} \times \mathbf{v}$ for all vectors $\mathbf{v}$.

We can introduce the linear operators $\texttt{sk}(\cdot)$ and $\texttt{sk}^{-1}(\cdot)$ to convert between a skew symmetric matrix and its corresponding cross-product vector:

$$\texttt{sk}\left(\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}, \quad \texttt{sk}^{-1}\left(\begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}\right) = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Applying $\texttt{sk}^{-1}$ to both sides of (3), we obtain:

$$\texttt{sk}^{-1}(Q^T C Q) = \begin{pmatrix} \lambda_2 + \lambda_3 & 0 & 0 \\ 0 & \lambda_1 + \lambda_3 & 0 \\ 0 & 0 & \lambda_1 + \lambda_2 \end{pmatrix} \texttt{sk}^{-1}(Q^T M Q) = (\operatorname{tr}(\Lambda) I - \Lambda)\texttt{sk}^{-1}(Q^T M Q).$$

Next, we apply the transformation formula $\texttt{sk}^{-1}(Q^T B Q) = Q^T \texttt{sk}^{-1}(B)$ (rotating, taking a cross product with $\mathbf{b}$, and then rotating back is the same thing as rotating $\mathbf{b}$ back and taking a cross product):

$$\begin{aligned} Q^T \texttt{sk}^{-1}(C) &= (\operatorname{tr}(\Lambda) I - \Lambda) Q^T \texttt{sk}^{-1}(M) \quad \iff \\ \texttt{sk}^{-1}(C) &= Q(\operatorname{tr}(\Lambda) I - \Lambda) Q^T \texttt{sk}^{-1}(M) = (\operatorname{tr}(S) I - S)\texttt{sk}^{-1}(M) \quad \iff \\ M &= \texttt{sk}\Big((\operatorname{tr}(S) I - S)^{-1} \texttt{sk}^{-1}(C)\Big). \end{aligned}$$

Finally, we plug in the expressions for C and M and solve for $\dot{R}$:

$$R^T \dot{R} = (\operatorname{tr}(S) I - S)^{-1} \texttt{sk}^{-1}(R^T \dot{A} - \dot{A}^T R) \quad \iff$$

$$\boxed{\dot{R} = R\, \texttt{sk}\Big(2(\operatorname{tr}(S) I - S)^{-1} \texttt{sk}^{-1}(R^T \dot{A})\Big),}$$

where we've extended the definition of $\texttt{sk}^{-1}(B)$ to non skew-symmetric matrices B by having it operate on the skew symmetric part $\frac{1}{2}\left(B - B^T\right)$.

Now that we know $\dot{R}$, it's easy to solve for $\dot{S}$ by going back to (2):

$$\boxed{\dot{S} = R^T \dot{A} - R^T \dot{R} S = R^T \Big(\dot{A} - \dot{R} S\Big).}$$