DERIVATIVE OF THE MATRIX INVERSE

ERIC PETERSON

Consider the normed vector space $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ of all linear operators of type signature $\mathbb{R}^n \to \mathbb{R}^n$. Among these, there is an interesting (open) subset U of invertible linear operators, as well as an interesting function

$$\chi\colon U \to U,$$
$$\chi(A) = A^{-1}$$

encoding the operation of matrix inversion. In this note, we will explore the derivative $D_A\chi$.

1. Manual computation

Manual computation is prohibitively complicated in general, but in the small setting of $n = 2$ we can work it all out for ourselves. Start by recalling the following explicit formula for χ:

$$\chi\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \frac{1}{\det A} \cdot \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

We consider the two pieces of the definition separately. The derivative of $(\det(-))^{-1}$ is expressed by

$$D_A(\det(-)^{-1}) = \frac{-1}{\det A^2} D_A \det = \frac{-1}{\det A^2}\begin{pmatrix} \frac{\partial \det A}{\partial a} & \frac{\partial \det A}{\partial b} & \frac{\partial \det A}{\partial c} & \frac{\partial \det A}{\partial d} \end{pmatrix} = \frac{-1}{\det A^2}\begin{pmatrix} d & -c & -b & a \end{pmatrix},$$

whereas the derivative of the matrix factor is itself expressed by a matrix of partial derivaties

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Finally, we use the product rule:

$$\begin{aligned}
D_A\chi &= \begin{pmatrix} d \\ -b \\ -c \\ a \end{pmatrix} \cdot \frac{-1}{\det A^2}\begin{pmatrix} d & -c & -b & a \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \cdot \frac{1}{\det A} \\
&= \begin{pmatrix} d^2 & -cd & -bd & ad \\ -bd & bc & b^2 & -ab \\ -cd & c^2 & bc & -ac \\ ad & -ac & -ab & a^2 \end{pmatrix} \cdot \frac{-1}{\det A^2} + \begin{pmatrix} 0 & 0 & 0 & -\det A \\ 0 & \det A & 0 & 0 \\ 0 & 0 & \det A & 0 \\ -\det A & 0 & 0 & 0 \end{pmatrix} \cdot \frac{-1}{\det A^2} \\
&= \begin{pmatrix} d^2 & -cd & -bd & bc \\ -bd & ad & b^2 & -ab \\ -cd & c^2 & ad & -ac \\ bc & -ac & -ab & a^2 \end{pmatrix} \cdot \frac{-1}{\det A^2}.
\end{aligned}$$

2. Slick computation

We now consider the defining property of the inverse of an operator:

$$A \cdot A^{-1} = I.$$

We can consider both sides of this equation as functions of A: the right-hand side is the constant function at I, and the left-hand side is the composite of several operations: duplicate A, invert one of the factors, and multiply them together. Since these agree as functions, their derivatives must also be equal. The derivative of the constant function is zero. The derivative of the composite is ripe for the chain rule:

$$D_A(\mu \circ (1, \chi))(H) = D_{(A,A^{-1})}\mu \circ D_A(1, \chi)(H) = D_{(A,A^{-1})}\mu \circ (H, (D_A\chi)(H)).$$

The multiplication function μ is bilinear, so Thayer's homework assignment gives us a product rule:

$$D_{(A,A^{-1})}\mu \circ (H, (D_A\chi)(H)) = A \circ (D_A\chi)(H) + H \circ A^{-1}.$$

Setting these two computations equal to each other (i.e., setting this one equal to zero) and solving for $D_A\chi$ gives

$$(D_A\chi)(H) = -A^{-1} \circ H \circ A^{-1}.$$

You should note the similarity between this answer and the classical calculation

$$P_1^{1/x@a}(a + h) = \frac{1}{a} + \frac{-1}{a^2} \cdot h$$

of the linear approximation to the function $1/x$ at a.

3. Comparison

The answer from the slick section admits explicit expansion, with a little tolerance. We take the answer and expand it as a function of H:

$$\begin{aligned} -A^{-1}HA^{-1} &= \frac{-1}{\det A^2} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} h & j \\ i & k \end{pmatrix} \cdot \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \\ &= \frac{-1}{\det A^2} \cdot \begin{pmatrix} d(dh - ci) - b(dj - ck) & a(dj - ck) - c(dh - ci) \\ d(ai - bh) - b(ak - bj) & a(ak - bj) - c(ai - bh) \end{pmatrix}. \end{aligned}$$

Notice that the pieces of the right-hand side are all linear in the components of H, so that this can be rearranged into a 4×4 matrix expressing the new components of H in terms of the old:

$$= \frac{-1}{\det A^2} \begin{pmatrix} d^2 & -cd & -bd & bc \\ -bd & da & b^2 & -ab \\ -cd & c^2 & ad & -ac \\ bc & -ac & -ab & a^2 \end{pmatrix} \cdot \begin{pmatrix} h \\ i \\ j \\ k \end{pmatrix}.$$

The observation, then, is that this matrix agrees with the one from the first section (but the formula $-A^{-1}HA^{-1}$ holds for all n and is easier to memorize besides).[1]

[1] I might have messed up a blanket minus sign in here. Sorry.