Christian B. Mendl, Richard M. Milbradt

Exercise 13.1 (Derivative of matrix inversion)
Our goal is to "differentiate" the inversion of a matrix, $A \mapsto A^{-1}$, for non-singular $A \in \mathbb{R}^{n\times n}$. In *forward mode* differentiation, we imagine that the entries of A depend on some real parameter x, and use the notation $\dot{A} = \frac{\mathrm{d}}{\mathrm{d}x} A(x)$. The inverse matrix is referred to as $C = A^{-1}$ in the following.

(a) Observe that $CA = I$ is constant (independent of x), such that $\frac{\mathrm{d}}{\mathrm{d}x}(CA) = 0$. Combine this equation with the product rule for differentiation to derive that

$$\dot{C} = -C\dot{A}C. \tag{1}$$

To explain *reverse mode* differentiation, let us consider a vector-valued differentiable function $f : \mathbb{R} \to \mathbb{R}^n$, $y = f(x)$. Imagine that, possibly after a sequence of calculations, y determines the scalar output of a "cost" function $\mathcal{L}$. We use the notation $\overline{y} = \frac{\partial \mathcal{L}}{\partial y}$ for the gradient vector of $\mathcal{L}$ with respect to y. Indirectly $\mathcal{L}$ also depends on x through f. By the chain rule,

$$\overline{x} = \frac{\partial \mathcal{L}}{\partial x} = \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial y_i} \frac{\partial y_i}{\partial x} = \langle \overline{y}, \nabla f(x) \rangle . \tag{2}$$

(b) In the context of matrix inversion, we identify y with C (interpreted as vector) and x with one of the entries of A. Use that $\langle \mathrm{vec}(B), \mathrm{vec}(A) \rangle = \mathrm{tr}[B^T A]$ for real matrices A, B and Eqs. (**??**), (**??**), to derive that

$$\overline{A} = -C^T\, \overline{C}\, C^T.$$

Hint: First compute $\overline{a_{ij}}$, with a_{ij} one of the entries of A and the other entries assumed constant, using Eq. (**??**) for $\partial C / \partial a_{ij}$.

Solution

(a) The product rule works for matrices as for scalar quantities, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}x}(AB) = \dot{A}B + A\dot{B}$$

for any matrices A and B of compatible dimensions, where each product is the usual matrix-matrix multiplication. (One can verify this statement by expanding AB in its matrix entries.) In our case,

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}(CA) = \dot{C}A + C\dot{A},$$

thus

$$\dot{C}A = -C\dot{A}, \qquad \dot{C} = -C\dot{A}A^{-1} = -C\dot{A}C.$$

As a remark, note that Eq. (**??**) generalizes the inversion of scalar quantities, $a \mapsto \frac{1}{a}$, with derivative $-\frac{1}{a^2}$.

(b) Expressing Eq. (**??**) in terms of matrices, we obtain

$$\overline{a_{ij}} = \mathrm{tr}\Big[\overline{C}^T \frac{\partial C}{\partial a_{ij}}\Big] \overset{\text{Eq. (??)}}{=} -\,\mathrm{tr}\Big[\overline{C}^T C \frac{\partial A}{\partial a_{ij}} C\Big] = -\,\mathrm{tr}\Big[C\, \overline{C}^T C E_{ij}\Big] = -\Big(C\, \overline{C}^T C\Big)_{ji}$$

where $E_{ij} = \partial A / \partial a_{ij}$ is the "unit" matrix with a single nonzero entry 1 at index (i, j). For the third equal sign we have also used the cyclic invariance of the trace: $\mathrm{tr}[AB] = \mathrm{tr}[BA]$. Assembling $\overline{a_{ij}}$ for all i, j in the matrix $\overline{A}$, we conclude that $\overline{A}$ is the transpose of $-C\, \overline{C}^T C$:

$$\overline{A} = -\Big(C\, \overline{C}^T C\Big)^T = -C^T\, \overline{C}\, C^T.$$