Matrix derivatives

P. Chainais (see Bishop, p. 697-698)

The following formulas are often useful to solve optimization problems implying matrix formulations.

The basic formulas are

$$\left(\frac{\partial x}{\partial \mathbf{a}}\right)_i = \frac{\partial x}{\partial a_i} \tag{1}$$

$$\left(\frac{\partial \mathbf{a}}{\partial \mathbf{b}}\right)_{ij} = \frac{\partial a_i}{\partial b_j} \tag{2}$$

which lead to

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{a}^T\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T\mathbf{a}) = \mathbf{a}. \tag{3}$$

If $\mathbf{A}$ and $\mathbf{B}$ depend on some variable x

$$\frac{\partial}{\partial x}(\mathbf{AB}) = \frac{\partial \mathbf{A}}{\partial x}\mathbf{B} + \mathbf{A}\frac{\partial \mathbf{B}}{\partial x} \tag{4}$$

The derivative of the inverse of a matrix is

$$\frac{\partial}{\partial x}(\mathbf{A}^{-1}) = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial x}\mathbf{A}^{-1} \tag{5}$$

Denoting $|\mathbf{A}| = \det\mathbf{A}$:

$$\frac{\partial}{\partial x}(\ln|\mathbf{A}|) = \mathrm{Tr}\left(\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial x}\right) \tag{6}$$

The trace of a matrix product often appears, so that we will need

$$\frac{\partial}{\partial A_{ij}}\mathrm{Tr}(\mathbf{AB}) = B_{ji} \tag{7}$$

which yields in compact notations:

$$\frac{\partial}{\partial \mathbf{A}}\mathrm{Tr}(\mathbf{AB}) = \mathbf{B}^T \tag{8}$$

With this notation we have the following properties:

$$\frac{\partial}{\partial \mathbf{A}}\mathrm{Tr}(\mathbf{A}^T\mathbf{B}) = \mathbf{B} \tag{9}$$

$$\frac{\partial}{\partial \mathbf{A}}\mathrm{Tr}(\mathbf{A}) = \mathbf{I} \tag{10}$$

$$\frac{\partial}{\partial \mathbf{A}}\mathrm{Tr}(\mathbf{ABA}^T) = \mathbf{A}(\mathbf{B}+\mathbf{B}^T) \tag{11}$$

$$\frac{\partial}{\partial \mathbf{A}}\mathrm{Tr}(\mathbf{ABA}^T\mathbf{C}) = \mathbf{C}^T\mathbf{AB}^T + \mathbf{CAB} \tag{12}$$

We also have

$$\frac{\partial}{\partial \mathbf{A}}\ln|\mathbf{A}| = (\mathbf{A}^{-1})^T \tag{13}$$