Gradients of Matrices

Useful identities:

$$egin{aligned} & rac{\partial}{\partial X}f(X)^{ op} = \left(rac{\partial f(X)}{\partial X}
ight)^{ op} \ & rac{\partial}{\partial X} ext{tr}(f(X)) = ext{tr}\left(rac{\partial f(X)}{\partial X}
ight) \ & rac{\partial}{\partial X} ext{det}(f(X)) = ext{det}(f(X)) ext{tr}\left(f(X)^{-1}rac{\partial f(X)}{\partial X}
ight) \end{aligned}$$

$$\frac{\partial}{\partial X}\det(f(X)) = \det(f(X))\mathrm{tr}\left(f(X)^{-1}\frac{\partial f(X)}{\partial X}\right)$$

(5.101)

(5.100)

Note:

Trace of a $D \times D \times E \times F$ tensor $\in \mathbb{R}^{E \times F}$

(5.99)

$$egin{aligned} rac{\partial}{\partial X}f(X)^{-1} &= -f(X)^{-1}rac{\partial f(X)}{\partial X}f(X)^{-1} \ rac{\partial a^ op X^{-1}b}{\partial X} &= -(X^{-1})^ op ab^ op (X^{-1})^ op \ rac{\partial x^ op a}{\partial X} &= a^ op \end{aligned}$$

$$rac{\partial a^ op X^{-1}b}{\partial X} = -(X^{-1})^ op ab^ op (X^{-1})^ op$$

$$rac{\partial x^{ op}a}{\partial x} = a^{ op}$$
 $rac{\partial a^{ op}x}{\partial x} = a^{ op}$

$$\frac{\partial x}{\partial x} = a^{\top}$$
 $\frac{\partial a^{\top} X b}{\partial a^{\top} X} = a b^{\top}$

$$rac{\partial a^ op X}{\partial X} = ab^ op \ rac{\partial x^ op Bx}{\partial x} = x^ op (B+B^ op)$$

$$rac{\partial x^{+}Bx}{\partial x} = x^{\top}(B+B^{\top})$$

$$\frac{\partial x^{\top} B x}{\partial x} = x^{\top} (B + B^{\top})$$
 (5.10)
$$\frac{\partial}{\partial s} (x - As)^{\top} W (x - As) = -2(x - As)^{\top} W A \text{ for symmetric } W$$

(5.108)

(5.107)

For x_{ij} , $-\mathbf{f}(\mathbf{X})^{-1} \frac{\partial \mathbf{f}(\mathbf{X})}{\partial x_{ij}} \mathbf{f}(\mathbf{X})^{-1} \in \mathbb{R}^{p \times p}$

 $\rightarrow -\mathbf{f}(\mathbf{X})^{-1} \frac{\partial \mathbf{f}(\mathbf{X})}{\partial \mathbf{X}} \mathbf{f}(\mathbf{X})^{-1} \in \mathbb{R}^{p \times p \times m \times n}$

12

(5.106)

 $\frac{\partial \mathbf{f}(\mathbf{X})}{\partial \mathbf{X}} \in \mathbb{R}^{p \times p \times m \times n}$

(5.105)

Right term: $\mathbf{f}(\mathbf{X})^{-1} \in \mathbb{R}^{p \times p}$

(5.104)

(5.103)

Left term: $\frac{\partial}{\partial \mathbf{X}} \mathbf{f}(\mathbf{X})^{-1} \in \mathbb{R}^{p \times p \times m \times n}$

 $(5.102) \quad \mathbf{Y} \in \mathbb{R}^{m \times n}, \ \mathbf{f}(\mathbf{X}) \in \mathbb{R}^{p \times p}$

$$rac{\partial x^ op B x}{\partial x} = x^ op (B + B^ op)$$

$$rac{\partial x^ op B x}{\partial x} = x^ op (B + B^ op)$$

$$\dfrac{\partial x^{ op} B x}{\partial x} = x^{ op} (B + B^{ op})$$

$$rac{\partial x^{ op}Bx}{\partial x} = x^{ op}(B+B^{ op})$$

$$\dfrac{x^ op B x}{\partial x} = x^ op (B + B^ op)$$

$$rac{\partial x^ op B x}{\partial x} = x^ op (B + B^ op)$$

$$\frac{\partial x^{ op}}{\partial x} = x^{ op}(B+B^{ op})$$

$$\frac{\partial x^{ op} B x}{\partial x} = x^{ op} (B + B^{ op})$$

$$rac{\partial a^ op X b}{\partial X} = a b^ op$$

$$\frac{\partial \bar{x}}{\partial x} = a^{\top}$$

$$\frac{^{ op}Xb}{\mathbf{x}}=ab^{^{ op}}$$

$$\frac{Xb}{X} = ab^{\top}$$

$$\frac{\partial C}{\partial x} = ab^{-1}$$

$$\frac{\Delta b}{X} = ab^{ op}$$

$$\frac{\mathbf{A} \, o}{c} = a b^{ op}$$

$$\frac{\Delta}{\zeta} = ab^{ op}$$

$$\frac{\partial X}{\partial X} = ab^{ op}$$

$$\frac{1}{x} = a^{ op}$$

$$\frac{^{ op}x}{}=a^{ op}$$

$$\frac{a}{\partial x} = a^{\top}$$

$$\frac{1}{2} = a^{-1}$$

$$a_{-\tilde{a}^{ op}}$$

$$\overline{}^{\!\scriptscriptstyle o} = {m a}^{\scriptscriptstyle op}$$

Gradients of Matrices

Clarification of some identities:

$$egin{aligned} rac{\partial}{\partial X}f(X)^{-1} &= -f(X)^{-1}rac{\partial f(X)}{\partial X}f(X)^{-1} \end{aligned}$$

$$\bullet \quad 0 = \frac{\partial I}{\partial X} = \frac{\partial}{\partial X} \left(f(X)^{-1} f(X) \right) = \left(\frac{\partial}{\partial X} f(X)^{-1} \right) f(X) + f(X)^{-1} \left(\frac{\partial}{\partial X} f(X) \right)$$

$$rac{\partial}{\partial s}(x-As)^{ op}W(x-As) = -2(x-As)^{ op}WA$$
 for symmetric W

Let
$$\mathbf{z} = \mathbf{x} - \mathbf{A}\mathbf{s}$$
, and use $\frac{\partial x^{\top}Bx}{\partial x} = x^{\top}(B + B^{\top})$:

$$\frac{\partial \mathbf{z}^{\mathsf{T}} \mathbf{W} \mathbf{z}}{\partial \mathbf{z}} = \mathbf{z}^{\mathsf{T}} (\mathbf{W} + \mathbf{W}^{\mathsf{T}}) = 2(\mathbf{x} - \mathbf{A}\mathbf{s})^{\mathsf{T}} \mathbf{W}$$

Then compute
$$\frac{\partial \mathbf{z}^T \mathbf{w} \mathbf{z}}{\partial \mathbf{z}} \frac{\partial \mathbf{z}}{\partial \mathbf{s}} = 2(\mathbf{x} - \mathbf{A}\mathbf{s})^T \mathbf{W} \cdot (-\mathbf{A})$$