15 Derivatives of a vector and its functions

15.1 Scalar function of a scalar

$$f(x+\Delta x) = f(x) + \frac{df}{dx}\Delta x$$

The increment in the function is the product of the derivative (gradient of f) and increment in the argument.

15.2 Scalar function of a vector

A function of N variables where the N variables are the components of a vector.

$$\begin{aligned}
\overline{c} &= \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} \\
f(\overline{c}) &= f(c_1, c_2, \ldots, c_N) \\
\overline{\nabla}_c f &= \begin{bmatrix} \frac{\partial f}{\partial c_1} \\ \frac{\partial f}{\partial c_2} \\ \vdots \\ \frac{\partial f}{\partial c_N} \end{bmatrix} \\
f(\overline{c}+\overline{\Delta c}) &= f(\overline{c}) + \overline{\nabla}_c^T f(\overline{c}) \cdot \overline{\Delta c} \\
\overline{g}(\overline{c}) &= \begin{bmatrix} g_1(c_1, c_2, \ldots, c_N) \\ g_2(c_1, c_2, \ldots, c_N) \\ \vdots \\ g_N(c_1, c_2, \ldots, c_N) \end{bmatrix}
\end{aligned}$$

$\overline{\nabla}_c f$ is the gradient of $f(\overline{c})$. The increment in f is the *dot product* of the gradient of f and the increment vector $\overline{\Delta c}$.

15.3 Vector function of a vector

$$\begin{aligned}
\nabla_c \overline{g} &= \begin{bmatrix} \frac{\partial g_1}{\partial c_1} & \frac{\partial g_2}{\partial c_1} & \cdots & \frac{\partial g_N}{\partial c_1} \\ \frac{\partial g_1}{\partial c_2} & \frac{\partial g_2}{\partial c_2} & \cdots & \frac{\partial g_N}{\partial c_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial c_N} & \frac{\partial g_2}{\partial c_N} & \cdots & \frac{\partial g_N}{\partial c_N} \end{bmatrix} \\
&= \begin{bmatrix} \overline{\nabla}_c g_1 \overline{\nabla}_c g_2 \ldots \overline{\nabla}_c g_N \end{bmatrix} \\
\overline{g}(\overline{c}+\overline{\Delta c}) &= \overline{g}(\overline{c}) + \nabla_c^T \overline{g}\, \overline{\Delta c}
\end{aligned}$$

$\nabla_c \overline{g}$ is the Jacobian of $\overline{g}(\overline{c})$. The columns of the Jacobian are the gradients of each component of the vector $\overline{g}$. The increment in each component of $\overline{g}$ is the dot product of the gradient of each component of $\overline{g}$ and the increment vector $\overline{\Delta c}$.

15.4 Vector times a constant

$$
\begin{aligned}
A &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix} \\
A\overline{c} &= \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + \ldots + a_{1N}c_N \\ a_{21}c_1 + a_{22}c_2 + \ldots + a_{2N}c_N \\ \vdots \\ a_{N1}c_1 + a_{N2}c_2 + \ldots + a_{NN}c_N \end{bmatrix} \\
\overline{c}^T A &= \begin{bmatrix} a_{11}c_1 + a_{21}c_2 + \ldots + a_{N1}c_N \\ a_{12}c_1 + a_{22}c_2 + \ldots + a_{N2}c_N \\ \vdots \\ a_{1N}c_1 + a_{2N}c_2 + \ldots + a_{NN}c_N \end{bmatrix} \\
\nabla(A\overline{c}) &= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{N1} \\ a_{12} & a_{22} & \dots & a_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1N} & a_{2N} & \dots & a_{NN} \end{bmatrix} \\
&= A^T \\
\nabla(\overline{c}^T A) &= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{N1} \\ a_{12} & a_{22} & \dots & a_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1N} & a_{2N} & \dots & a_{NN} \end{bmatrix} \\
&= A^T
\end{aligned}
$$

15.5 Scalar function of a vector times a constant

$$
\begin{aligned}
f(\overline{c}) &= \overline{b}^T \overline{c} \\
&= \overline{c}^T \overline{b} \\
&= b_1 c_1 + b_2 c_2 + \ldots + b_N c_N \\
\overline{\nabla}_c f &= \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \\
&= \overline{b} \\
f(\overline{c})\,\overline{b} &= \begin{bmatrix} b_1 f(\overline{c}) \\ b_2 f(\overline{c}) \\ \vdots \\ b_N f(\overline{c}) \end{bmatrix} \\
\nabla_c \left(f(\overline{c})\,\overline{b} \right) &= \begin{bmatrix} b_1 \frac{\partial f}{\partial c_1} & b_2 \frac{\partial f}{\partial c_1} & \ldots & b_N \frac{\partial f}{\partial c_1} \\ b_1 \frac{\partial f}{\partial c_2} & b_2 \frac{\partial f}{\partial c_2} & \ldots & b_N \frac{\partial f}{\partial c_2} \\ \vdots & \vdots & \ddots & \vdots \\ b_1 \frac{\partial f}{\partial c_N} & b_2 \frac{\partial f}{\partial c_N} & \ldots & b_N \frac{\partial f}{\partial c_N} \end{bmatrix} \\
&= \overline{\nabla}_c f\,\overline{b}^T
\end{aligned}
$$

15.6 Vector function of a vector times a constant

$$
\begin{aligned}
A\overline{g}(\overline{c}) &= \begin{bmatrix} a_{11}g_1 + a_{12}g_2 + \ldots + a_{1N}g_N \\ a_{21}g_1 + a_{22}g_2 + \ldots + a_{2N}g_N \\ \vdots \\ a_{N1}g_1 + a_{N2}g_2 + \ldots + a_{NN}g_N \end{bmatrix} \\
\overline{\nabla}_c(A\overline{g}) &= \begin{bmatrix} a_{11}\frac{\partial g_1}{\partial c_1} + a_{12}\frac{\partial g_2}{\partial c_1} + \ldots + a_{1N}\frac{\partial g_N}{\partial c_1} \\ a_{21}\frac{\partial g_1}{\partial c_2} + a_{22}\frac{\partial g_2}{\partial c_2} + \ldots + a_{2N}\frac{\partial g_N}{\partial c_2} \\ \vdots \\ a_{N1}\frac{\partial g_1}{\partial c_N} + a_{N2}\frac{\partial g_2}{\partial c_N} + \ldots + a_{NN}\frac{\partial g_N}{\partial c_N} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial g_1}{\partial c_1} & \frac{\partial g_2}{\partial c_1} & \cdots & \frac{\partial g_N}{\partial c_1} \\ \frac{\partial g_1}{\partial c_2} & \frac{\partial g_2}{\partial c_2} & \cdots & \frac{\partial g_N}{\partial c_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial c_N} & \frac{\partial g_2}{\partial c_N} & \cdots & \frac{\partial g_N}{\partial c_N} \end{bmatrix} A^T \\
&= \nabla_c\overline{g}\, A^T
\end{aligned}
$$

15.7 Norm of a vector and related functions

$$\begin{aligned}
f(\overline{c}) &= \overline{c}^T\overline{c}\\
&= c_1^2 + c_2^2 + \ldots + c_N^2\\
\overline{\nabla}\left(\overline{c}^T\overline{c}\right) &= \begin{bmatrix} 2c_1 \\ 2c_2 \\ \vdots \\ 2c_N \end{bmatrix}\\
&= 2\overline{c}\\
f(\overline{c}) &= \overline{c}^T A\overline{c}\\
&= a_{11}c_1^2 + a_{12}c_1c_2 + \ldots + a_{1N}c_1c_N\\
&\quad + a_{21}c_2c_1 + a_{22}c_2^2 + \ldots + a_{2N}c_2c_N\\
&\quad + \ldots\\
&\quad + a_{N1}c_Nc_1 + a_{N2}c_Nc_2 + \ldots + a_{NN}c_N^2\\
\overline{\nabla}_c f &= \begin{bmatrix} 2a_{11}c_1 + (a_{12}+a_{21})c_2 + \ldots + (a_{1N}+a_{N1})c_N \\ (a_{21}+a_{12})c_1 + 2a_{22}c_2 + \ldots + (a_{2N}+a_{N2})c_N \\ \vdots \\ (a_{N1}+a_{1N})c_1 + (a_{N2}+a_{2N})c_2 + \ldots + 2a_{NN}c_N \end{bmatrix}\\
&= \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + \ldots + a_{1N}c_N \\ a_{21}c_1 + a_{22}c_2 + \ldots + a_{2N}c_N \\ \vdots \\ a_{N1}c_1 + a_{N2}c_2 + \ldots + a_{NN}c_N \end{bmatrix} + \begin{bmatrix} a_{11}c_1 + a_{21}c_2 + \ldots + a_{N1}c_N \\ a_{12}c_1 + a_{22}c_2 + \ldots + a_{N2}c_N \\ \vdots \\ a_{1N}c_1 + a_{2N}c_2 + \ldots + a_{NN}c_N \end{bmatrix}\\
&= \left(A + A^T\right)\overline{c}
\end{aligned}$$

15.8 Least mean square fitting

Choose $\overline{c}$ such that the norm of the "error vector" $\overline{e} = A\overline{c} - \overline{b}$ is minimized. A is an $M \times N$ matrix and $\overline{b}$ is an $M \times 1$ vector.

$$\begin{aligned}
\overline{e} &= A\overline{c} - \overline{b}\\
\overline{e}^T\overline{e} &= \left(A\overline{c} - \overline{b}\right)^T\left(A\overline{c} - \overline{b}\right)\\
&= \left(\overline{c}^T A^T - \overline{b}^T\right)\left(A\overline{c} - \overline{b}\right)\\
&= \overline{c}^T A^T A\overline{c} - \overline{b}^T A\overline{c} - \overline{c}^T A^T\overline{b} - \overline{b}^T\overline{b}
\end{aligned}$$

Using the gradients of $\bar{c}^T A\bar{c}$, $\bar{b}^T\bar{c}$, and $\bar{c}^T\bar{b}$ derived above,

$$\begin{aligned}\nabla_c\left(\bar{e}^T\bar{e}\right) &= \left(A^TA+(A^TA)^T\right)\bar{c}-A^T\bar{b}-A^T\bar{b}\\ &= 2A^TA\bar{c}-2A^T\bar{b}\end{aligned}$$

Setting the gradient to zero for minimization, we get the least mean square fit equation

$$\bar{c} = \left(A^TA\right)^{-1}A^T\bar{b}$$

In case of an FIR equalizer, columns of A are delayed versions of the unit pulse response, and $\bar{c}=\left(A^TA\right)^{-1}A^T\bar{b}$ is the coefficient vector optimized for LMS error.

15.9 Norm of a vector function of a vector

$$\begin{aligned}f(\bar{c}) &= \bar{g}^T\bar{g}\\ &= g_1^2+g_2^2+\ldots+g_N^2\\ \overline{\nabla}_c f &= \begin{bmatrix}2g_1\frac{\partial g_1}{\partial c_1}+2g_2\frac{\partial g_2}{\partial c_1}+\ldots+2g_N\frac{\partial g_N}{\partial c_1}\\ 2g_1\frac{\partial g_1}{\partial c_2}+2g_2\frac{\partial g_2}{\partial c_2}+\ldots+2g_N\frac{\partial g_N}{\partial c_2}\\ \vdots\\ 2g_1\frac{\partial g_1}{\partial c_N}+2g_2\frac{\partial g_2}{\partial c_N}+\ldots+2g_N\frac{\partial g_N}{\partial c_N}\end{bmatrix}\\ &= 2\begin{bmatrix}\frac{\partial g_1}{\partial c_1} & \frac{\partial g_2}{\partial c_1} & \cdots & \frac{\partial g_N}{\partial c_1}\\ \frac{\partial g_1}{\partial c_2} & \frac{\partial g_2}{\partial c_2} & \cdots & \frac{\partial g_N}{\partial c_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial g_1}{\partial c_N} & \frac{\partial g_2}{\partial c_N} & \cdots & \frac{\partial g_N}{\partial c_N}\end{bmatrix}\bar{g}\\ &= 2\left(\nabla_c\bar{g}\right)\bar{g}\end{aligned}$$

15.10 Least mean square fitting

The result in the previous section can be used to find the gradient of $\left(A\bar{c}-\bar{b}\right)^T\left(A\bar{c}-\bar{b}\right)$ in a straightforward way. Setting $\bar{g}=A\bar{c}-\bar{b}$, $\overline{\nabla}_c(\bar{g}^T\bar{g})=2[\nabla_c(A\bar{c}-\bar{b})](A\bar{c}-\bar{b})=2A^T(A\bar{c}-\bar{b})$