Differentials

1 Scalar function of scalar variable

Let $f : \mathbb{R} \to \mathbb{R}$ be a scalar function of scalar variable. The idea of a differential for a function like that is pretty standard:

$$\delta f = \frac{df}{dx}\delta x.$$

It is often used to e.g. define a linear approximation to f at a point $\hat{x}$:

$$g(\hat{x}, x) = \frac{df}{dx}(\hat{x})(x - \hat{x}) + f(\hat{x})$$

I.e. $g(\hat{x}, x)$ is a very good approximation of $f(x)$ for x near $\hat{x}$. Now, I want to be a little more rigorous in defining the differential δf in a manner similar to this. That is, $\delta f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

$$\delta f(\hat{x}, \delta x) = \frac{df}{dx}(\hat{x})\delta x, \quad \hat{x} \in \mathbb{R}, \ \delta x \in \mathbb{R}.$$

With this definition

$$g(\hat{x}, x) = \delta f(\hat{x}, x - \hat{x}) + f(\hat{x})$$

2 More general functions

For functions defined over higher dimensional spaces, it can actually be preferable to work with differentials rather than derivatives because amongst other reasons, formulas tend to be more compact. For example, if I have a function $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$, then

$$\delta f : \mathbb{R}^{2\times 2} \times \mathbb{R}^{2\times 2} \to \mathbb{R}, \quad \delta f(\hat{\mathbf{F}}, \delta \mathbf{F}) = \frac{\partial f}{\partial F_{ij}}(\hat{\mathbf{F}})\delta F_{ij}, \quad \hat{\mathbf{F}}, \delta\mathbf{F} \in \mathbb{R}^{2\times 2}$$

Also for a matrix valued function $\mathbf{G} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$,

$$\delta \mathbf{G} : \mathbb{R}^{2\times 2} \times \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}, \quad \delta \mathbf{G}(\hat{\mathbf{F}}, \delta \mathbf{F}) = \frac{\partial \mathbf{G}}{\partial F_{ij}}(\hat{\mathbf{F}})\delta F_{ij}, \quad \hat{\mathbf{F}}, \delta\mathbf{F} \in \mathbb{R}^{2\times 2}$$

this can also be written as

$$\delta G_{kl}(\hat{\mathbf{F}}, \delta \mathbf{F}) = \frac{\partial G_{kl}}{\partial F_{ij}}(\hat{\mathbf{F}})\delta F_{ij}, \quad \hat{\mathbf{F}}, \delta\mathbf{F} \in \mathbb{R}^{2\times 2}$$

.

2.1 Example: determinant

Let's consider the example of $J : \mathbb{R}^{2\times 2} \to \mathbb{R}$, $J(\mathbf{F}) = \det(\mathbf{F}) = F_{11}F_{22} - F_{21}F_{12}$ and

$$\frac{\partial J}{\partial \mathbf{F}}(\hat{\mathbf{F}}) = J(\hat{\mathbf{F}})\hat{\mathbf{F}}^{-T} \text{ and } \delta J(\hat{\mathbf{F}}, \delta \mathbf{F}) = J(\hat{\mathbf{F}})\hat{F}_{ji}^{-1}\delta F_{ij}$$

2.2 Example: trace

Consider $\mathrm{Tr} : \mathbb{R}^{2\times 2} \to \mathbb{R}$, $\mathrm{Tr}(\mathbf{F}) = F_{ii}$ and

$$\frac{\partial \mathrm{Tr}}{\partial F_{kl}}(\mathbf{F}) = \delta_{kl}, \text{ and } \delta\mathrm{Tr}(\hat{\mathbf{F}}, \delta\mathbf{F}) = \delta_{kl}\delta F_{kl} = \delta F_{kk} = \delta F_{ii}.$$

I.e. $\delta\mathrm{Tr}(\hat{\mathbf{F}}, \delta\mathbf{F}) = \mathrm{Tr}(\delta\mathbf{F})$. In general, for any linear function $\mathbf{G}$, $\delta\mathbf{G}(\hat{\mathbf{F}}, \delta\mathbf{F}) = \mathbf{G}(\delta\mathbf{F})$.

2.3 Example: SVD

We can think of the SVD of a matrix as a function of the entries in the matrix:

$$\mathbf{F} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{U}(\mathbf{F})\mathbf{\Sigma}(\mathbf{F})\mathbf{V}^T(\mathbf{F})$$

where

$$\mathbf{U} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$$
$$\mathbf{\Sigma} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$$
$$\mathbf{V} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}.$$

Of course, the mapping is not bijective since $\mathbf{U}$ and $\mathbf{V}$ are always orthogonal and Σ is always diagonal. It is very convenient to work out the differentials of $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Sigma}$ (especially by comparison to the explicit derivative formulas). First, it can be verified that differentials satisfy a product rule. We can use this to show that since

$$\mathbf{U}^T(\mathbf{F})\mathbf{U}(\mathbf{F}) = \mathbf{I}, \quad \mathbf{V}^T(\mathbf{F})\mathbf{V}(\mathbf{F}) = \mathbf{I}$$

we can see

$$\delta\mathbf{U}^T\mathbf{U} + \mathbf{U}^T\delta\mathbf{U} = \mathbf{0}, \quad \delta\mathbf{V}^T\mathbf{V} + \mathbf{V}^T\delta\mathbf{V} = \mathbf{0}$$

and this implies that $\delta\mathbf{U}^T\mathbf{U}$ and $\delta\mathbf{V}^T\mathbf{V}$ are skew-symmetric. We can combine this with

$$\delta\mathbf{F} = \delta\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T + \mathbf{U}\delta\mathbf{\Sigma}\mathbf{V}^T + \mathbf{U}\mathbf{\Sigma}\delta\mathbf{V}^T$$

and therefore

$$\mathbf{U}^T\delta\mathbf{F}\mathbf{V} = \mathbf{U}^T\delta\mathbf{U}\mathbf{\Sigma} + \delta\mathbf{\Sigma} + \mathbf{\Sigma}\delta\mathbf{V}^T\mathbf{V}$$

to solve for $\delta\mathbf{\Sigma}$, $\delta\mathbf{V}$ and $\delta\mathbf{U}$ in terms of $\delta\mathbf{F}$, $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Sigma}$.