Integration By Substitution

Philip Schulz and Wilker Aziz
https://github.com/vitutorial/VITutorial

Multivariate calculus recap

Let $x \in \mathbb{R}^K$ and let $\mathcal{T} : \mathbb{R}^K \to \mathbb{R}^K$ be differentiable and invertible

- $y = \mathcal{T}(x)$
- $x = \mathcal{T}^{-1}(y)$

Jacobian

The Jacobian matrix $J_{\mathcal{T}}(x)$ of $\mathcal{T}$
assessed at x is the matrix of partial derivatives

$$J_{ij} = \frac{\partial y_i}{\partial x_j} = \frac{\partial \mathcal{T}(x)_i}{\partial x_j}$$

Jacobian

The Jacobian matrix $J_{\mathcal{T}}(x)$ of $\mathcal{T}$
assessed at x is the matrix of partial derivatives

$$J_{ij} = \frac{\partial y_i}{\partial x_j} = \frac{\partial \mathcal{T}(x)_i}{\partial x_j}$$

Inverse function theorem

$$J_{\mathcal{T}^{-1}}(y) = \left(J_{\mathcal{T}}(x)\right)^{-1}$$

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Scalar Case

$$\mathrm{d}y \frac{\mathrm{d}x}{\mathrm{d}x} =$$

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Scalar Case

$$\mathrm{d}y \frac{\mathrm{d}x}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x = \frac{\mathrm{d}\mathcal{T}(x)}{\mathrm{d}x}\mathrm{d}x$$

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Scalar Case

$$\mathrm{d}y\frac{\mathrm{d}x}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x = \frac{\mathrm{d}\mathcal{T}(x)}{\mathrm{d}x}\mathrm{d}x \implies \mathrm{d}x = \frac{\mathrm{d}x}{\mathrm{d}y}\mathrm{d}y$$

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Scalar Case

$$\mathrm{d}y\frac{\mathrm{d}x}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x = \frac{\mathrm{d}\mathcal{T}(x)}{\mathrm{d}x}\mathrm{d}x \implies \mathrm{d}x = \frac{\mathrm{d}x}{\mathrm{d}y}\mathrm{d}y$$

- $\mathrm{d}y/\mathrm{d}x$ scales the differential $\mathrm{d}x$ to match it to $\mathrm{d}y$
- if $\mathrm{d}y/\mathrm{d}x > 1$, $\mathcal{T}$ expands the area around x locally

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Differential (or inifinitesimal)

The **differential** $\mathrm{d}x$ of x refers to an *infinitely small* change in x

We can relate the differential $\mathrm{d}y$ of $y = \mathcal{T}(x)$ to $\mathrm{d}x$

Multivariate Case

$$\mathrm{d}y = \left|\det J_{\mathcal{T}}(x)\right| \mathrm{d}x$$

the absolute value absorbs the orientation

Integration by substitution

We can integrate a function $g(x)$ by substituting $x = \mathcal{T}^{-1}(y)$

$$\int g(x)\,\mathrm{d}x$$

Integration by substitution

We can integrate a function $g(x)$
by substituting $x = \mathcal{T}^{-1}(y)$

$$\int g(x)\mathrm{d}x = \int g(\underbrace{\mathcal{T}^{-1}(y)}_{x})\,\underbrace{|\det J_{\mathcal{T}^{-1}}(y)|\mathrm{d}y}_{\mathrm{d}x}$$

Integration by substitution

We can integrate a function $g(x)$
by substituting $x = \mathcal{T}^{-1}(y)$

$$\int g(x)\mathrm{d}x = \int g(\underbrace{\mathcal{T}^{-1}(y)}_{x})\,\underbrace{|\det J_{\mathcal{T}^{-1}}(y)|\mathrm{d}y}_{\mathrm{d}x}$$

and similarly for a function $h(y)$

$$\int h(y)\mathrm{d}y$$

Integration by substitution

We can integrate a function $g(x)$
by substituting $x = \mathcal{T}^{-1}(y)$

$$\int g(x)\mathrm{d}x = \int g(\underbrace{\mathcal{T}^{-1}(y)}_{x}) \underbrace{|\det J_{\mathcal{T}^{-1}}(y)|\mathrm{d}y}_{\mathrm{d}x}$$

and similarly for a function $h(y)$

$$\int h(y)\mathrm{d}y = \int h(\mathcal{T}(x))|\det J_{\mathcal{T}}(x)|\mathrm{d}x$$

Change of density

Let X take on values in $\mathbb{R}^K$ with density $p_X(x)$

Change of density

Let X take on values in $\mathbb{R}^K$ with density $p_X(x)$
and recall that $y = \mathcal{T}(x)$ and $x = \mathcal{T}^{-1}(y)$

Change of density

Let X take on values in $\mathbb{R}^K$ with density $p_X(x)$
and recall that $y = \mathcal{T}(x)$ and $x = \mathcal{T}^{-1}(y)$

Then $\mathcal{T}$ induces a density $p_Y(y)$ expressed as

$$p_Y(y) = p_X(\mathcal{T}^{-1}(y))|\det J_{\mathcal{T}^{-1}}(y)|$$

Change of density

Let X take on values in $\mathbb{R}^K$ with density $p_X(x)$ and recall that $y = \mathcal{T}(x)$ and $x = \mathcal{T}^{-1}(y)$

Then $\mathcal{T}$ induces a density $p_Y(y)$ expressed as

$$p_Y(y) = p_X(\mathcal{T}^{-1}(y))|\det J_{\mathcal{T}^{-1}}(y)|$$

and then it follows that

$$p_X(x) = p_Y(\mathcal{T}(x))|\det J_{\mathcal{T}}(x)|$$