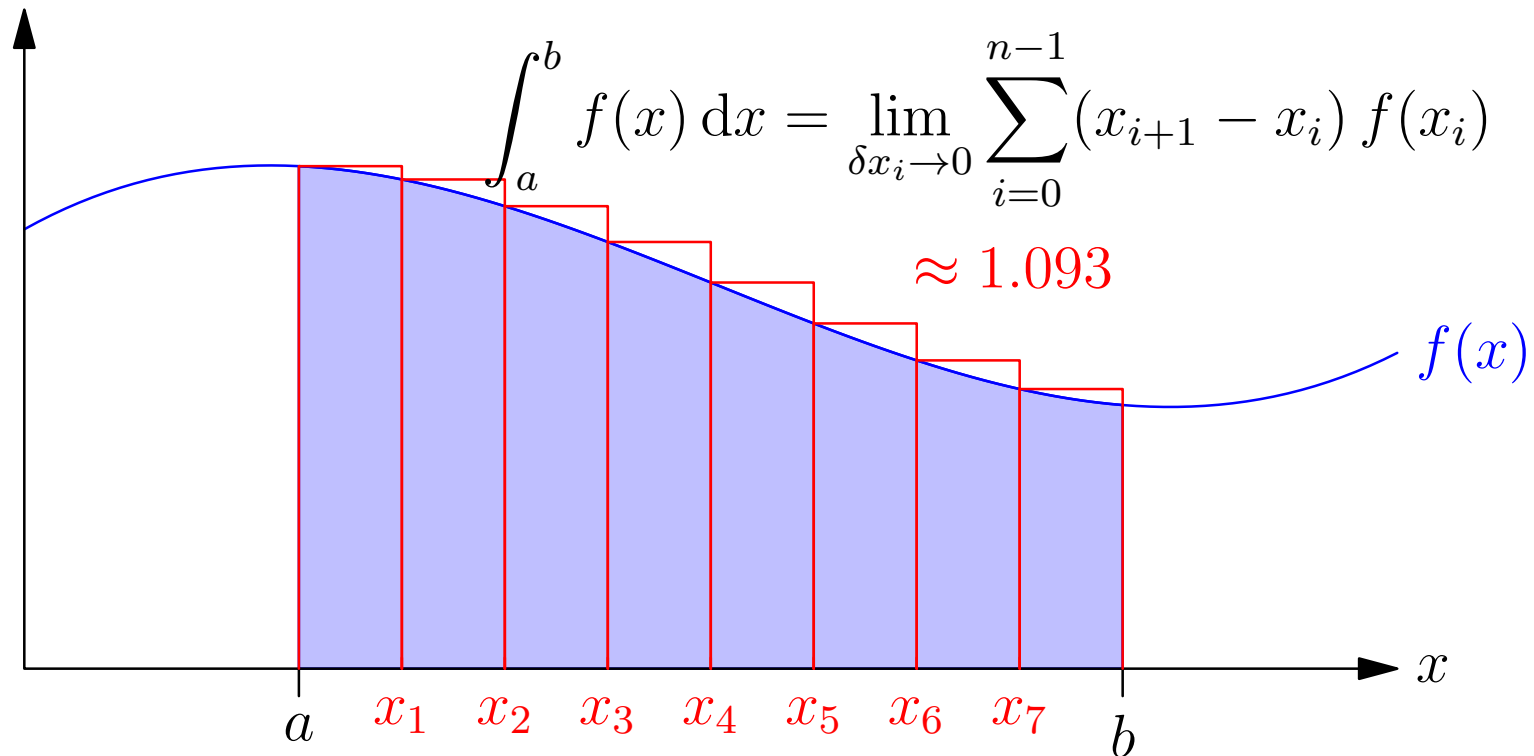


Advanced Machine Learning

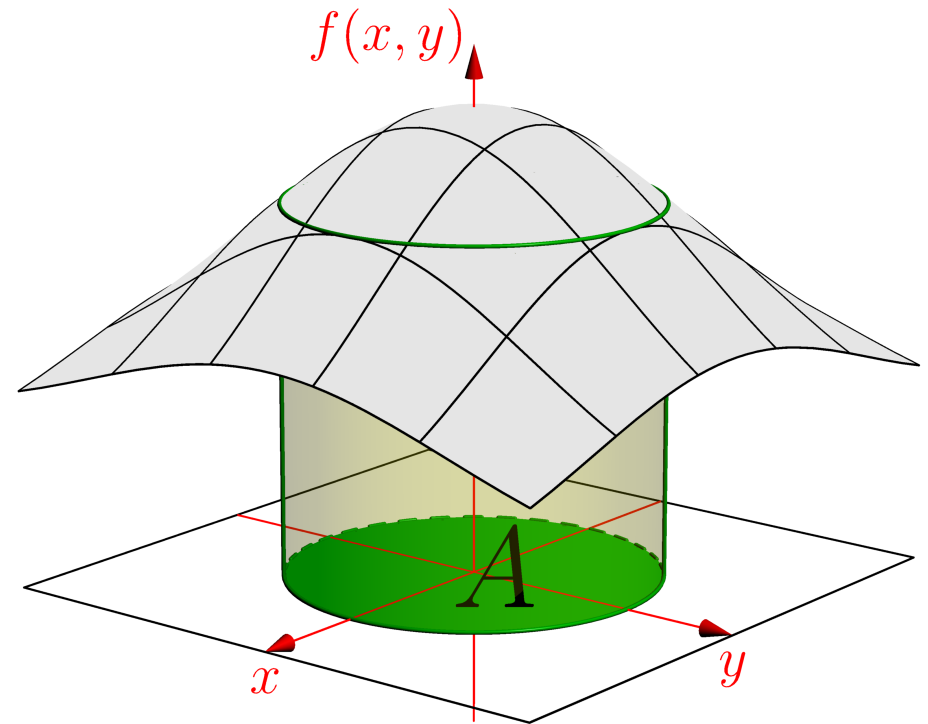
Integral Calculus



Riemann Integration, integration by parts, gaussian integrals

Outline

1. **Defining Integrals**
2. Doing Integrals
3. Gaussian Integrals

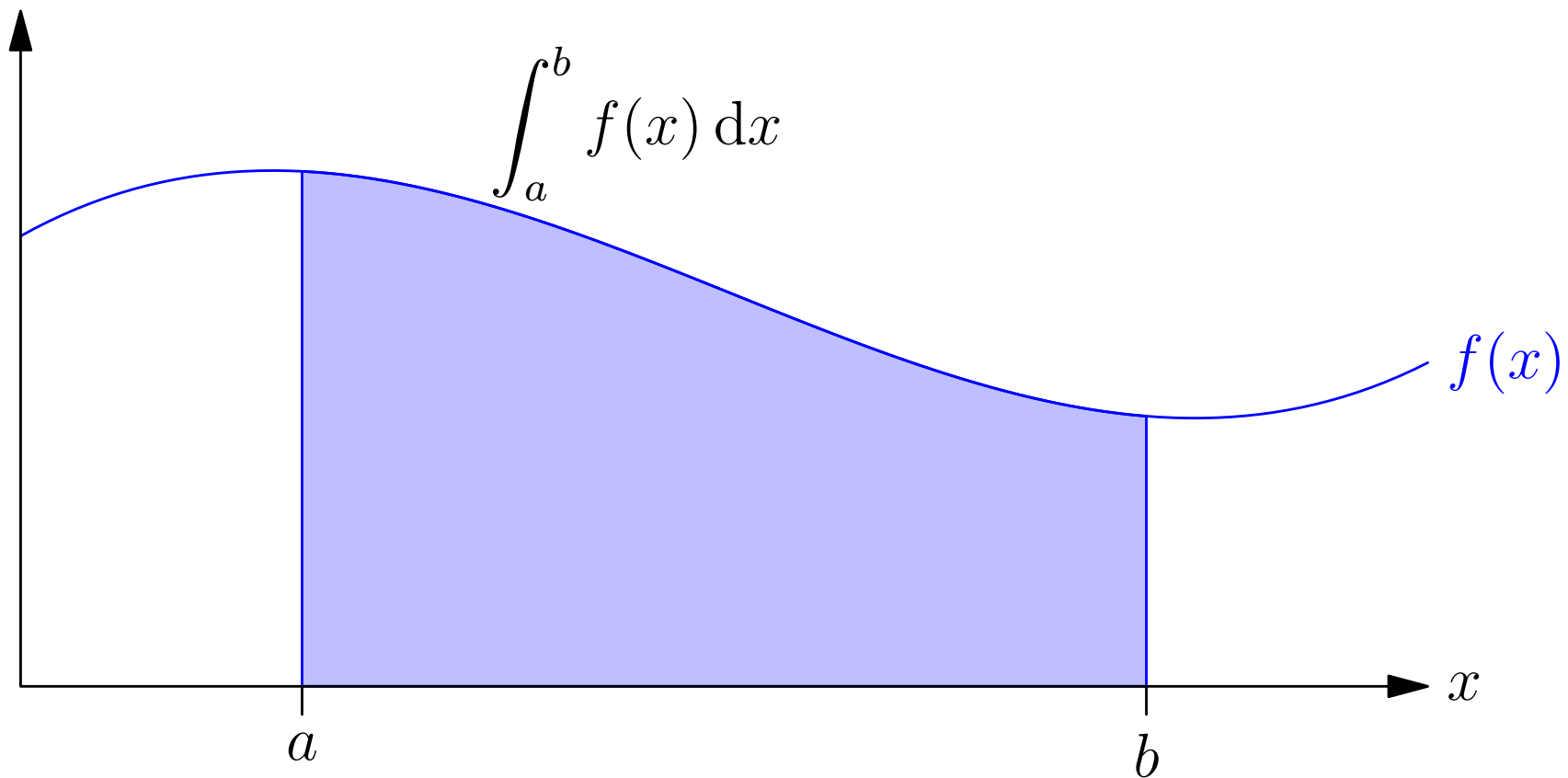


Riemann Integral

- Integrals represent area beneath a curve

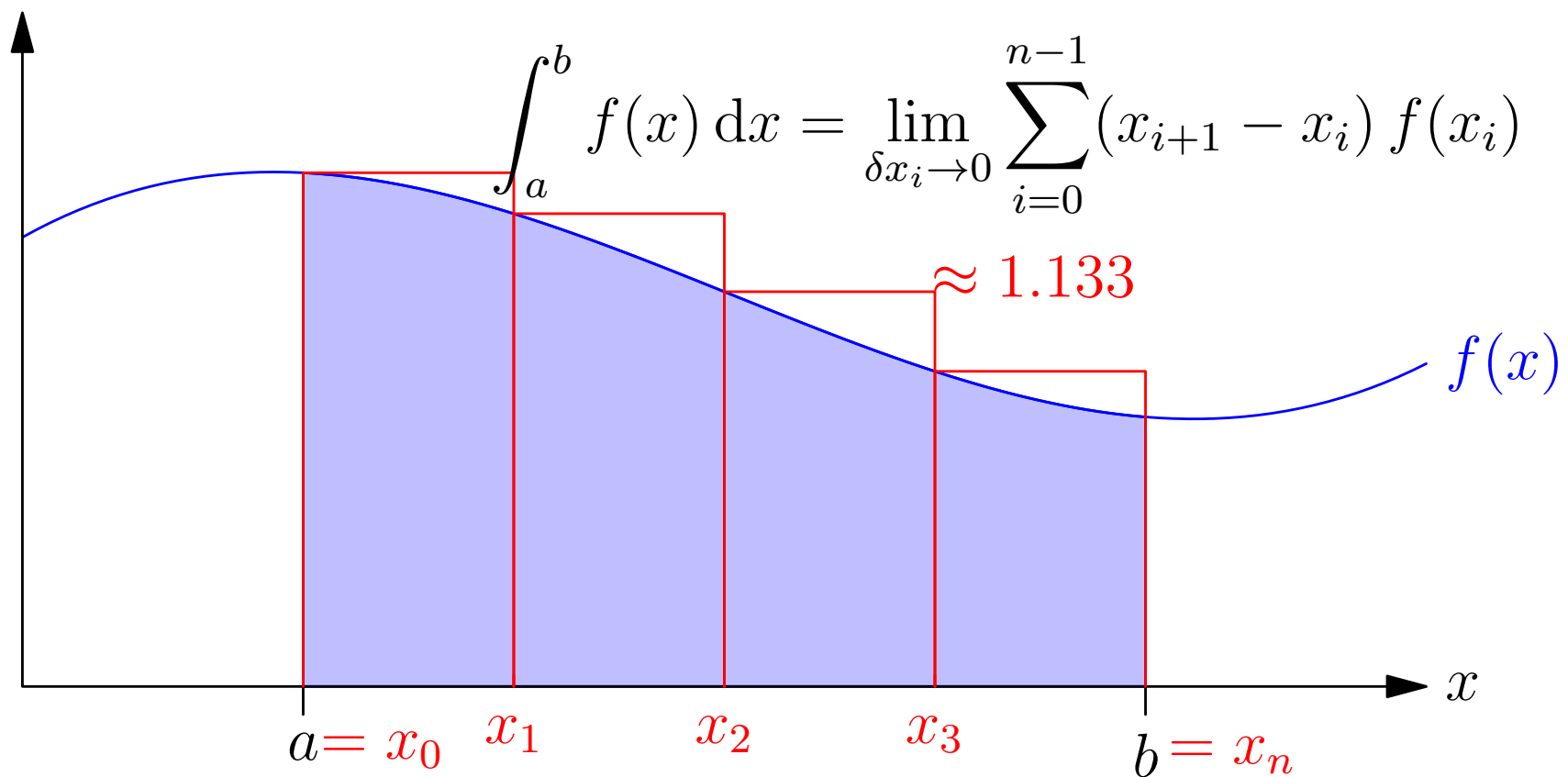
Riemann Integral

- Integrals represent area beneath a curve



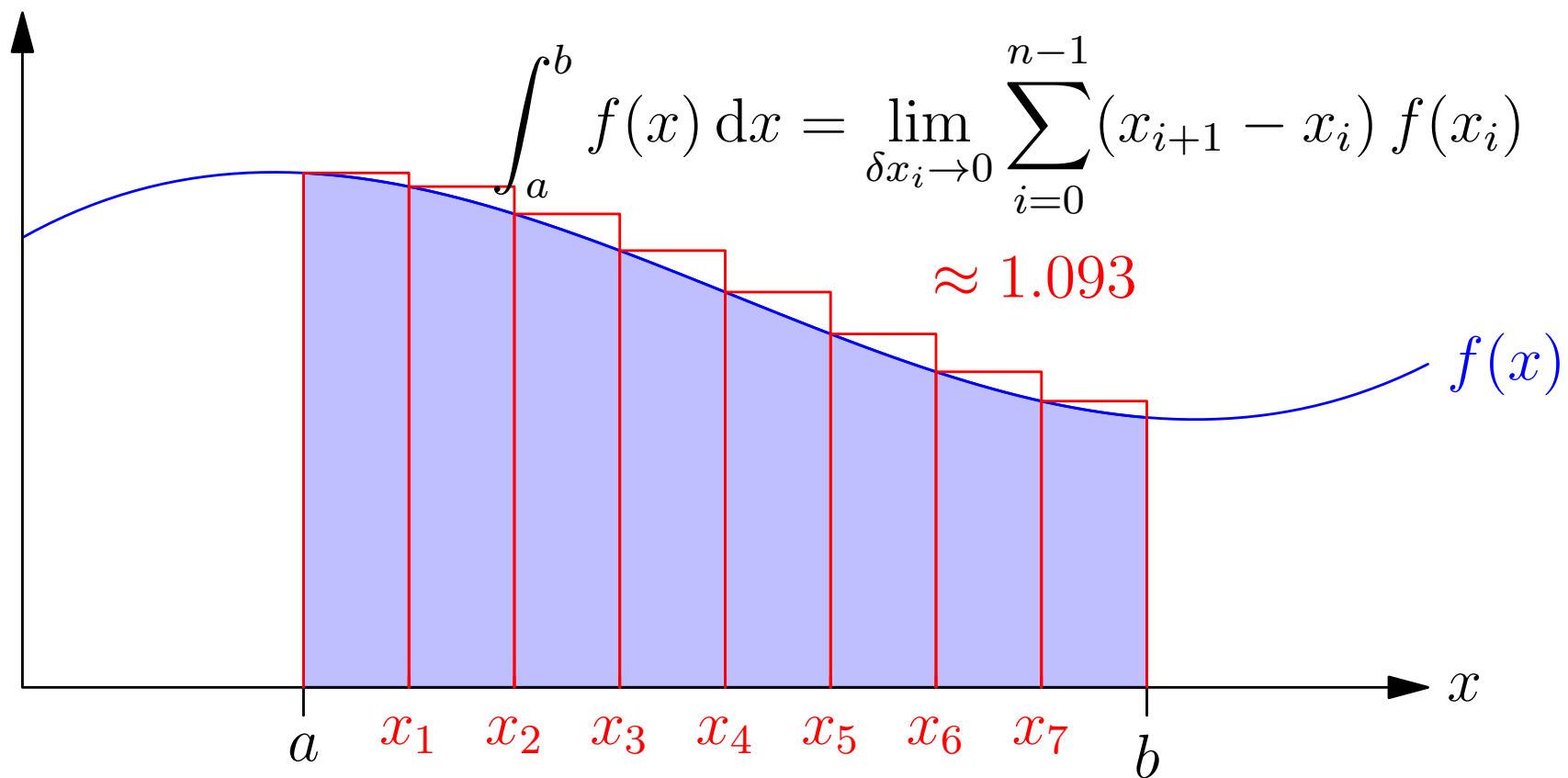
Riemann Integral

- Integrals represent area beneath a curve



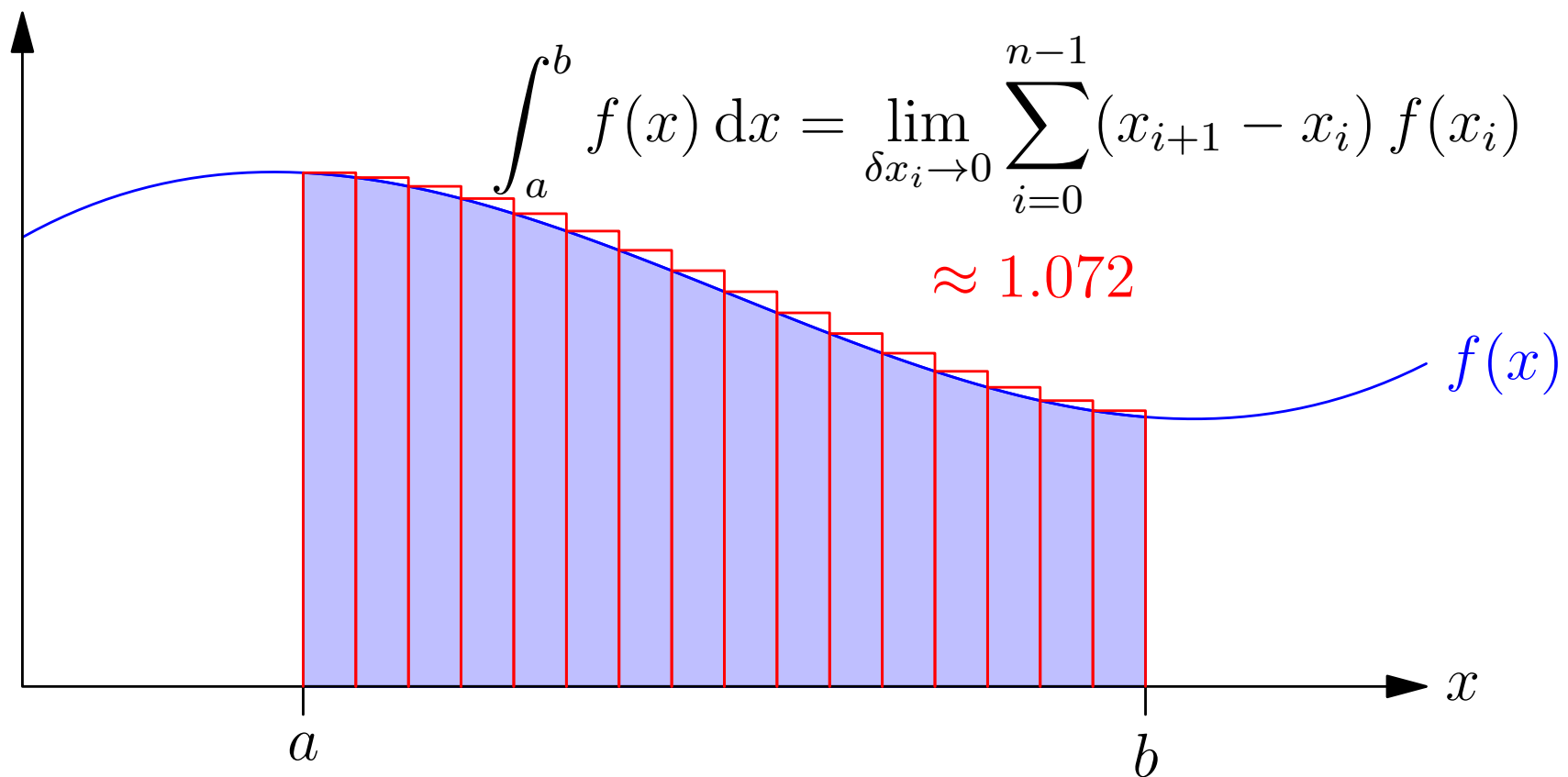
Riemann Integral

- Integrals represent area beneath a curve



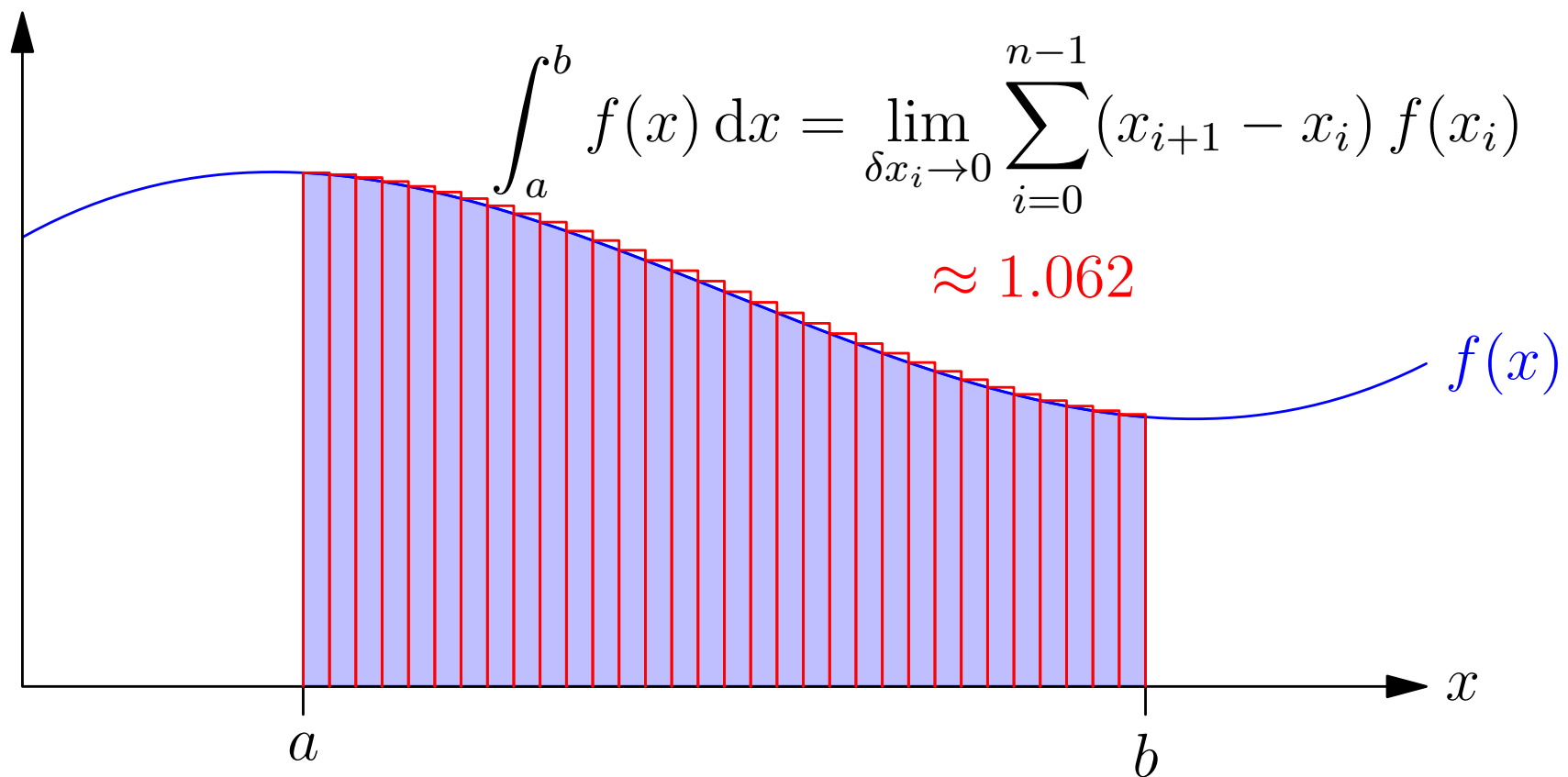
Riemann Integral

- Integrals represent area beneath a curve



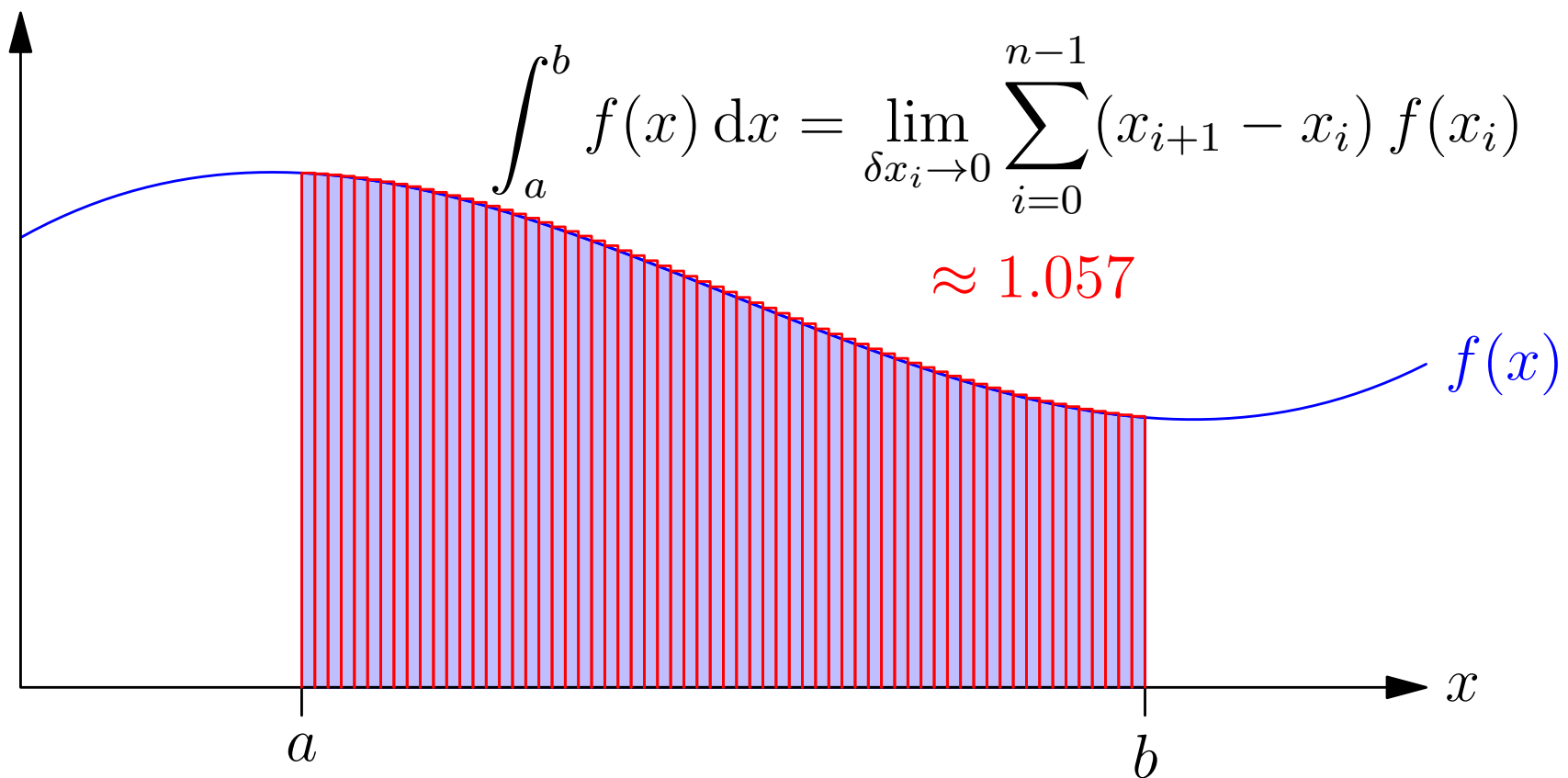
Riemann Integral

- Integrals represent area beneath a curve



Riemann Integral

- Integrals represent area beneath a curve



Linearity of Integration

- Integration is a linear operator

$$\int_a^b (r f(x) + s g(x)) dx = \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) (r f(x_i) + s g(x_i))$$

Linearity of Integration

- Integration is a linear operator

$$\begin{aligned}\int_a^b (r f(x) + s g(x)) dx &= \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) (r f(x_i) + s g(x_i)) \\ &= \lim_{\delta x_i \rightarrow 0} \left(\sum_{i=0}^{n-1} (x_{i+1} - x_i) r f(x_i) + \sum_{i=0}^{n-1} (x_{i+1} - x_i) s g(x_i) \right)\end{aligned}$$

Linearity of Integration

- Integration is a linear operator

$$\begin{aligned}\int_a^b (r f(x) + s g(x)) dx &= \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) (r f(x_i) + s g(x_i)) \\&= \lim_{\delta x_i \rightarrow 0} \left(\sum_{i=0}^{n-1} (x_{i+1} - x_i) r f(x_i) + \sum_{i=0}^{n-1} (x_{i+1} - x_i) s g(x_i) \right) \\&= \lim_{\delta x_i \rightarrow 0} \left(r \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(x_i) + s \sum_{i=0}^{n-1} (x_{i+1} - x_i) g(x_i) \right)\end{aligned}$$

Linearity of Integration

- Integration is a linear operator

$$\begin{aligned}\int_a^b (r f(x) + s g(x)) dx &= \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) (r f(x_i) + s g(x_i)) \\&= \lim_{\delta x_i \rightarrow 0} \left(\sum_{i=0}^{n-1} (x_{i+1} - x_i) r f(x_i) + \sum_{i=0}^{n-1} (x_{i+1} - x_i) s g(x_i) \right) \\&= \lim_{\delta x_i \rightarrow 0} \left(r \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(x_i) + s \sum_{i=0}^{n-1} (x_{i+1} - x_i) g(x_i) \right) \\&= r \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(x_i) + s \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) g(x_i)\end{aligned}$$

Linearity of Integration

- Integration is a linear operator

$$\begin{aligned}\int_a^b (r f(x) + s g(x)) dx &= \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) (r f(x_i) + s g(x_i)) \\&= \lim_{\delta x_i \rightarrow 0} \left(\sum_{i=0}^{n-1} (x_{i+1} - x_i) r f(x_i) + \sum_{i=0}^{n-1} (x_{i+1} - x_i) s g(x_i) \right) \\&= \lim_{\delta x_i \rightarrow 0} \left(r \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(x_i) + s \sum_{i=0}^{n-1} (x_{i+1} - x_i) g(x_i) \right) \\&= r \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(x_i) + s \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) g(x_i) \\&= r \int_a^b f(x) dx + s \int_a^b g(x) dx\end{aligned}$$

Fundamental Law of Calculus

- Let

$$I(a, x) = \int_a^x f(z) dz = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i)$$

- Now for small δx

$$I(a, x + \delta x) = \int_a^{x+\delta x} f(z) dz = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i) + \delta x f(x)$$

- Thus

$$\frac{dI(a, x)}{dx} = \lim_{\delta x \rightarrow 0} \frac{I(x + \delta x) - I(x)}{\delta x}$$

Fundamental Law of Calculus

- Let

$$I(a, x) = \int_a^x f(z) \mathrm{d}z = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i)$$

- Now for small δx

$$I(a, x + \delta x) = \int_a^{x+\delta x} f(z) \mathrm{d}z = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i) + \delta x f(x)$$

- Thus

$$\frac{\mathrm{d}I(a, x)}{\mathrm{d}x} = \lim_{\delta x \rightarrow 0} \frac{I(x + \delta x) - I(x)}{\delta x}$$

Fundamental Law of Calculus

- Let

$$I(a, x) = \int_a^x f(z) dz = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i)$$

- Now for small δx

$$I(a, x + \delta x) = \int_a^{x+\delta x} f(z) dz = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i) + \delta x f(x)$$

- Thus

$$\frac{dI(a, x)}{dx} = \lim_{\delta x \rightarrow 0} \frac{I(x + \delta x) - I(x)}{\delta x} = \lim_{\delta x \rightarrow 0} \frac{\delta x f(x)}{\delta x}$$

Fundamental Law of Calculus

- Let

$$I(a, x) = \int_a^x f(z) dz = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i)$$

- Now for small δx

$$I(a, x + \delta x) = \int_a^{x+\delta x} f(z) dz = \lim_{\delta z_i \rightarrow 0} \sum_{i=0}^{n-1} (z_{i+1} - z_i) f(z_i) + \delta x f(x)$$

- Thus

$$\frac{dI(a, x)}{dx} = \lim_{\delta x \rightarrow 0} \frac{I(x + \delta x) - I(x)}{\delta x} = \lim_{\delta x \rightarrow 0} \frac{\delta x f(x)}{\delta x} = f(x)$$

The Other Way Around

- Consider

$$\int_a^b \frac{df(x)}{dx} dx = \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx$$

The Other Way Around

- Consider

$$\begin{aligned}\int_a^b \frac{df(x)}{dx} dx &= \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx \\ &= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}\end{aligned}$$

The Other Way Around

- Consider

$$\begin{aligned}\int_a^b \frac{df(x)}{dx} dx &= \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx \\ &= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\ &= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (f(x_{i+1}) - f(x_i))\end{aligned}$$

The Other Way Around

- Consider

$$\begin{aligned}\int_a^b \frac{df(x)}{dx} dx &= \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (f(x_{i+1}) - f(x_i)) \\&= (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + (f(x_3) - f(x_2)) + \cdots \\&\quad + (f(x_{n-1}) - f(x_{n-2})) + (f(x_n) - f(x_{n-1}))\end{aligned}$$

The Other Way Around

- Consider

$$\begin{aligned}\int_a^b \frac{df(x)}{dx} dx &= \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (f(x_{i+1}) - f(x_i)) \\&= (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + (f(x_3) - f(x_2)) + \cdots \\&\quad + (f(x_{n-1}) - f(x_{n-2})) + (f(x_n) - f(x_{n-1})) \\&= f(x_n) - f(x_0)\end{aligned}$$

The Other Way Around

- Consider

$$\begin{aligned}\int_a^b \frac{df(x)}{dx} dx &= \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (f(x_{i+1}) - f(x_i)) \\&= (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + (f(x_3) - f(x_2)) + \cdots \\&\quad + (f(x_{n-1}) - f(x_{n-2})) + (f(x_n) - f(x_{n-1})) \\&= f(x_n) - f(x_0) = f(b) - f(a)\end{aligned}$$

The Other Way Around

- Consider

$$\begin{aligned}\int_a^b \frac{df(x)}{dx} dx &= \int_a^b \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x} dx \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (x_{i+1} - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\&= \lim_{x_{i+1} - x_i \rightarrow 0} \sum_{i=0}^{n-1} (f(x_{i+1}) - f(x_i)) \\&= (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + (f(x_3) - f(x_2)) + \dots \\&\quad + (f(x_{n-1}) - f(x_{n-2})) + (f(x_n) - f(x_{n-1})) \\&= f(x_n) - f(x_0) = f(b) - f(a)\end{aligned}$$

- We can think of integration as an **anti-derivative** it undoes differentiation

Indefinite Integrals

- So far we have considered **definite integrals** where we integrate between two points (a and b)
- However, when think about integration as an anti-derivative, it is useful to think of a function $F(x) = \int f(x)dx$
- So that $F'(x) = f(x)$
- However the function $F(x)$, $F(x) + 1$, $F(x) + \pi$, etc. all have the same derivative so $F(x)$ is only defined up to an additive constant
- Note that the definite integral is given by

$$\int_a^b f(x)dx = F(b) - F(a)$$

Indefinite Integrals

- So far we have considered **definite integrals** where we integrate between two points (a and b)
- However, when think about integration as an anti-derivative, it is useful to think of a function $F(x) = \int f(x)dx$
- So that $F'(x) = f(x)$
- However the function $F(x)$, $F(x) + 1$, $F(x) + \pi$, etc. all have the same derivative so $F(x)$ is only defined up to an additive constant
- Note that the definite integral is given by

$$\int_a^b f(x)dx = F(b) - F(a)$$

Indefinite Integrals

- So far we have considered **definite integrals** where we integrate between two points (a and b)
- However, when think about integration as an anti-derivative, it is useful to think of a function $F(x) = \int f(x)dx$
- So that $F'(x) = f(x)$
- However the function $F(x)$, $F(x) + 1$, $F(x) + \pi$, etc. all have the same derivative so $F(x)$ is only defined up to an additive constant
- Note that the definite integral is given by

$$\int_a^b f(x)dx = F(b) - F(a)$$

Indefinite Integrals

- So far we have considered **definite integrals** where we integrate between two points (a and b)
- However, when think about integration as an anti-derivative, it is useful to think of a function $F(x) = \int f(x)dx$
- So that $F'(x) = f(x)$
- However the function $F(x)$, $F(x) + 1$, $F(x) + \pi$, etc. all have the same derivative so $F(x)$ is only defined up to an additive constant
- Note that the definite integral is given by

$$\int_a^b f(x)dx = F(b) - F(a)$$

Indefinite Integrals

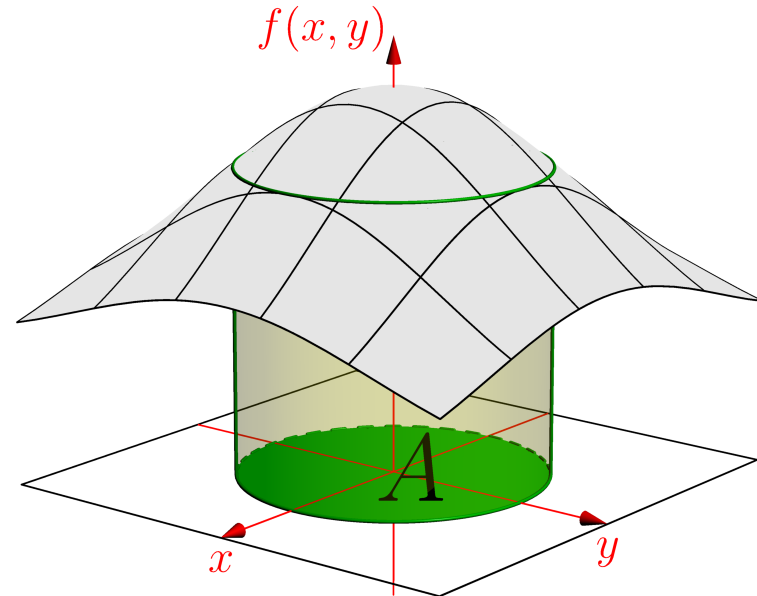
- So far we have considered **definite integrals** where we integrate between two points (a and b)
- However, when think about integration as an anti-derivative, it is useful to think of a function $F(x) = \int f(x)dx$
- So that $F'(x) = f(x)$
- However the function $F(x)$, $F(x) + 1$, $F(x) + \pi$, etc. all have the same derivative so $F(x)$ is only defined up to an additive constant
- Note that the definite integral is given by

$$\int_a^b f(x)dx = F(b) - F(a)$$

Multiple Integrals

- For functions involving many independent variables (e.g. $f(x,y)$, $f(x,y,z)$, $f(\mathbf{x})$) we can integrate over multiple dimensions
- For example

$$\iint_A f(x,y) dx dy$$



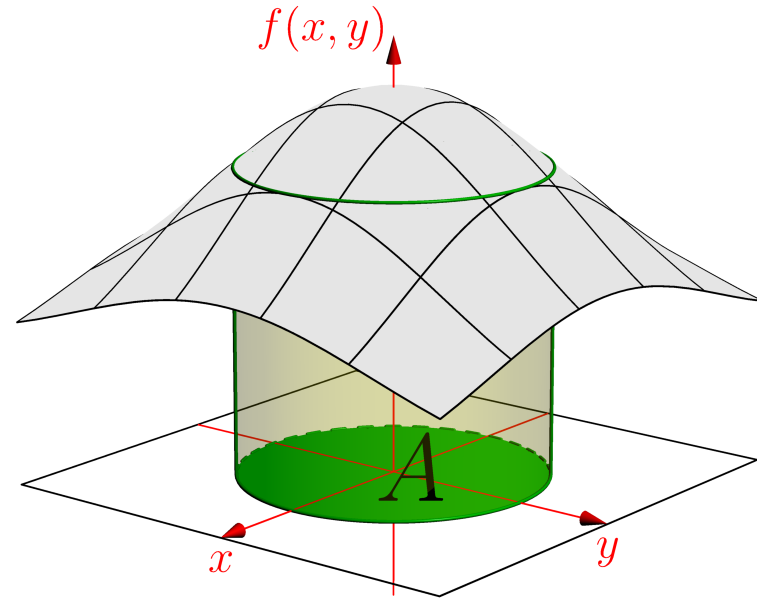
- It gets tedious writing multiple integral signs and I tend to write just one

$$\int \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n$$

Multiple Integrals

- For functions involving many independent variables (e.g. $f(x,y)$, $f(x,y,z)$, $f(\mathbf{x})$) we can integrate over multiple dimensions
- For example

$$\iint_A f(x,y) dx dy$$



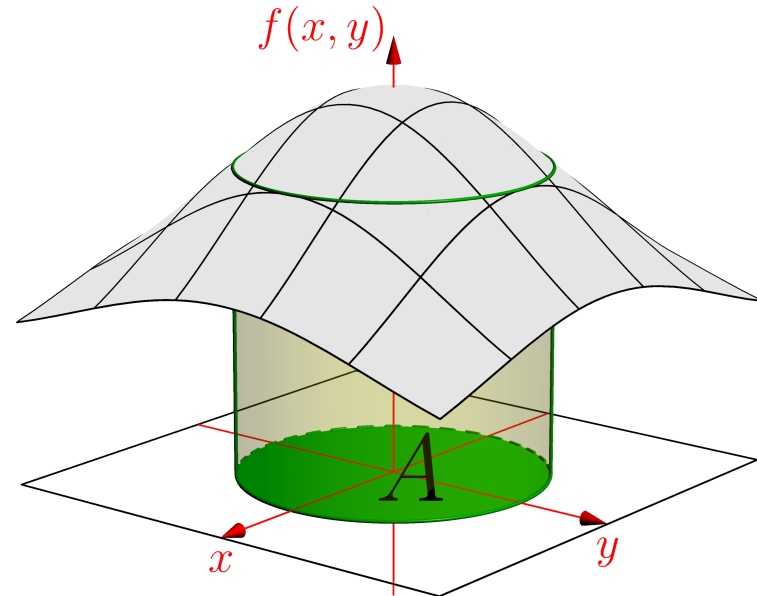
- It gets tedious writing multiple integral signs and I tend to write just one

$$\int \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n$$

Multiple Integrals

- For functions involving many independent variables (e.g. $f(x,y)$, $f(x,y,z)$, $f(\mathbf{x})$) we can integrate over multiple dimensions
- For example

$$\iint_A f(x,y) dx dy$$



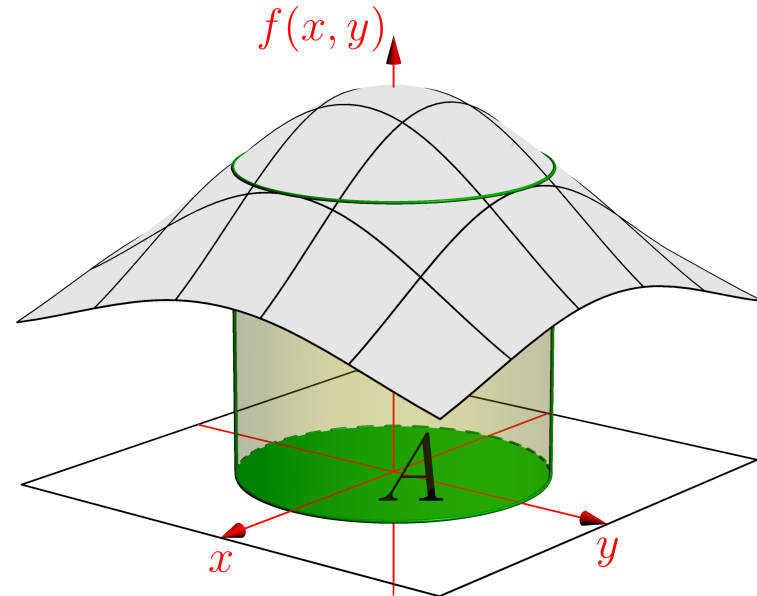
- It gets tedious writing multiple integral signs and I tend to write just one

$$\int \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n$$

Multiple Integrals

- For functions involving many independent variables (e.g. $f(x,y)$, $f(x,y,z)$, $f(\mathbf{x})$) we can integrate over multiple dimensions
- For example

$$\iint_A f(x,y) dx dy$$

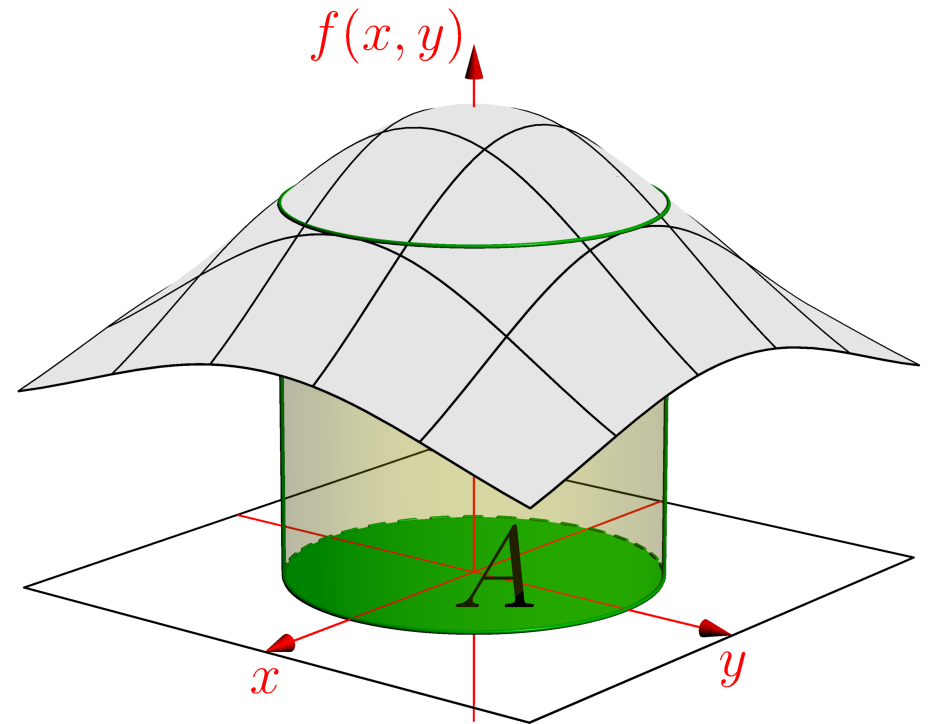


- It gets tedious writing multiple integral signs and I tend to write just one

$$\int \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = \int f(\mathbf{x}) d\mathbf{x}$$

Outline

1. Defining Integrals
2. **Doing Integrals**
3. Gaussian Integrals



Performing Integration

- A key method for performing integrals is through knowledge of the anti-derivative
- If we know $F'(x) = f(x)$ then $F(x) + c = \int f(x) dx$
- E.g. we know that $dx^n/dx = nx^{n-1}$ therefore

$$\int x^{n-1} dx = \frac{1}{n} \int \frac{dx^n}{dx} dx$$

and

$$\int_a^b x^{n-1} dx = \frac{b^n}{n} - \frac{a^n}{n}$$

Performing Integration

- A key method for performing integrals is through knowledge of the anti-derivative
- If we know $F'(x) = f(x)$ then $F(x) + c = \int f(x) dx$
- E.g. we know that $dx^n/dx = nx^{n-1}$ therefore

$$\int x^{n-1} dx = \frac{1}{n} \int \frac{dx^n}{dx} dx$$

and

$$\int_a^b x^{n-1} dx = \frac{b^n}{n} - \frac{a^n}{n}$$

Performing Integration

- A key method for performing integrals is through knowledge of the anti-derivative
- If we know $F'(x) = f(x)$ then $F(x) + c = \int f(x) dx$
- E.g. we know that $dx^n/dx = nx^{n-1}$ therefore

$$\int x^{n-1} dx = \frac{1}{n} \int \frac{dx^n}{dx} dx$$

and

$$\int_a^b x^{n-1} dx = \frac{b^n}{n} - \frac{a^n}{n}$$

Performing Integration

- A key method for performing integrals is through knowledge of the anti-derivative
- If we know $F'(x) = f(x)$ then $F(x) + c = \int f(x) dx$
- E.g. we know that $dx^n/dx = nx^{n-1}$ therefore

$$\int x^{n-1} dx = \frac{1}{n} \int \frac{dx^n}{dx} dx = \frac{x^n}{n}$$

and

$$\int_a^b x^{n-1} dx = \frac{b^n}{n} - \frac{a^n}{n}$$

Performing Integration

- A key method for performing integrals is through knowledge of the anti-derivative
- If we know $F'(x) = f(x)$ then $F(x) + c = \int f(x) dx$
- E.g. we know that $dx^n/dx = nx^{n-1}$ therefore

$$\int x^{n-1} dx = \frac{1}{n} \int \frac{dx^n}{dx} dx = \frac{x^n}{n} + c$$

and

$$\int_a^b x^{n-1} dx = \frac{b^n}{n} - \frac{a^n}{n}$$

Performing Integration

- A key method for performing integrals is through knowledge of the anti-derivative
- If we know $F'(x) = f(x)$ then $F(x) + c = \int f(x) dx$
- E.g. we know that $dx^n/dx = nx^{n-1}$ therefore

$$\int x^{n-1} dx = \frac{1}{n} \int \frac{dx^n}{dx} dx = \frac{x^n}{n} + c$$

and

$$\int_a^b x^{n-1} dx = \frac{b^n}{n} - \frac{a^n}{n}$$

Is Integration Straightforward?

- We saw due to the product and chain rules that we can differentiate almost anything

- Products and compositions

$$\int f(x)g(x)dx = ?$$

$$\int f(g(x))dx = ?$$

- Unfortunately, unlike differentiation we don't have a small parameter we can expand in
- In general integration is hard

Is Integration Straightforward?

- We saw due to the product and chain rules that we can differentiate almost anything. **Given integration is the anti-derivative can we integrate anything?**
- Products and compositions

$$\int f(x)g(x)dx = ? \qquad \int f(g(x))dx = ?$$

- Unfortunately, unlike differentiation we don't have a small parameter we can expand in
- In general integration is hard

Is Integration Straightforward?

- We saw due to the product and chain rules that we can differentiate almost anything. Given integration is the anti-derivative can we integrate anything?
- Products and compositions

$$\int f(x)g(x)dx = ?$$

$$\int f(g(x))dx = ?$$

- Unfortunately, unlike differentiation we don't have a small parameter we can expand in
- In general integration is hard

Is Integration Straightforward?

- We saw due to the product and chain rules that we can differentiate almost anything. Given integration is the anti-derivative can we integrate anything?
- Products and compositions

$$\int f(x)g(x)dx = ?$$

$$\int f(g(x))dx = ?$$

- Unfortunately, unlike differentiation we don't have a small parameter we can expand in
- In general integration is hard

Is Integration Straightforward?

- We saw due to the product and chain rules that we can differentiate almost anything. Given integration is the anti-derivative can we integrate anything?
- Products and compositions

$$\int f(x)g(x)dx = ? \qquad \int f(g(x))dx = ?$$

- Unfortunately, unlike differentiation we don't have a small parameter we can expand in
- In general integration is hard

Integration by Parts

- Recall the product rule $\frac{d}{dx} f(x)g(x) = \frac{d}{dx} f(x)g(x) + f(x)\frac{d}{dx} g(x)$
- Integrating we get

$$\int_a^b \frac{d}{dx} f(x)g(x) dx = \int_a^b \frac{d}{dx} f(x)g(x) dx + \int_a^b f(x)\frac{d}{dx} g(x) dx$$

- Unfortunately we get two integrals

Integration by Parts

- Recall the product rule $\frac{d}{dx} f(x)g(x) = \frac{d}{dx} f(x)g(x) + f(x)\frac{d}{dx} g(x)$
- Integrating we get

$$\int_a^b \frac{d}{dx} f(x)g(x) dx = \int_a^b \frac{d}{dx} f(x)g(x) dx + \int_a^b f(x)\frac{d}{dx} g(x) dx$$

- Unfortunately we get two integrals

Integration by Parts

- Recall the product rule $\frac{d}{dx} f(x)g(x) = \frac{d}{dx} f(x)g(x) + f(x)\frac{d}{dx} g(x)$
- Integrating we get

$$\begin{aligned}\int_a^b \frac{d}{dx} f(x)g(x) dx &= \int_a^b \frac{d}{dx} f(x)g(x) dx + \int_a^b f(x)\frac{d}{dx} g(x) dx \\ &= [f(x)g(x)]_a^b\end{aligned}$$

- Unfortunately we get two integrals

Integration by Parts

- Recall the product rule $\frac{d}{dx}(f(x)g(x)) = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx}$
- Integrating we get

$$\begin{aligned}\int_a^b \frac{d}{dx}(f(x)g(x)) dx &= \int_a^b \frac{df(x)}{dx}g(x) dx + \int_a^b f(x)\frac{dg(x)}{dx} dx \\ &= [f(x)g(x)]_a^b = f(b)g(b) - f(a)g(a)\end{aligned}$$

- Unfortunately we get two integrals

Integration by Parts

- Recall the product rule $\frac{d}{dx} f(x)g(x) = \frac{d}{dx} f(x)g(x) + f(x)\frac{d}{dx} g(x)$
- Integrating we get

$$\begin{aligned}\int_a^b \frac{d}{dx} f(x)g(x) dx &= \int_a^b \frac{d}{dx} f(x)g(x) dx + \int_a^b f(x)\frac{d}{dx} g(x) dx \\ &= [f(x)g(x)]_a^b = f(b)g(b) - f(a)g(a)\end{aligned}$$

- Unfortunately we get two integrals

Integration by Parts

- Recall the product rule $\frac{d}{dx} f(x)g(x) = \frac{d}{dx} f(x)g(x) + f(x)\frac{d}{dx} g(x)$
- Integrating we get

$$\begin{aligned}\int_a^b \frac{d}{dx} f(x)g(x) dx &= \int_a^b \frac{d}{dx} f(x)g(x) dx + \int_a^b f(x)\frac{d}{dx} g(x) dx \\ &= [f(x)g(x)]_a^b = f(b)g(b) - f(a)g(a)\end{aligned}$$

- Unfortunately we get two integrals, but we can turn this around

$$\int_a^b f(x)\frac{d}{dx} g(x) dx = [f(x)g(x)]_a^b - \int_a^b \frac{d}{dx} f(x)g(x) dx$$

Integration by Parts

- Recall the product rule $\frac{d}{dx} f(x)g(x) = \frac{d}{dx} f(x)g(x) + f(x)\frac{d}{dx} g(x)$
- Integrating we get

$$\begin{aligned}\int_a^b \frac{d}{dx} f(x)g(x) dx &= \int_a^b \frac{d}{dx} f(x)g(x) dx + \int_a^b f(x)\frac{d}{dx} g(x) dx \\ &= [f(x)g(x)]_a^b = f(b)g(b) - f(a)g(a)\end{aligned}$$

- Unfortunately we get two integrals, but we can turn this around

$$\int_a^b f(x)\frac{d}{dx} g(x) dx = [f(x)g(x)]_a^b - \int_a^b \frac{d}{dx} f(x)g(x) dx$$

whether this is helpful depends on $f(x)$ and $g(x)$

Example of Integration by Parts

- Consider

$$\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$$

Example of Integration by Parts

- Consider

$$\Pi(z) = \int_0^{\infty} x^z e^{-x} dx = \int_0^{\infty} x^z \frac{d(-e^{-x})}{dx} dx$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= \left[x^z (-e^{-x}) \right]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx\end{aligned}$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx\end{aligned}$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx = z \int_0^\infty x^{z-1} e^{-x} dx\end{aligned}$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Gamma(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx = z \int_0^\infty x^{z-1} e^{-x} dx = z\Gamma(z-1)\end{aligned}$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx = z \int_0^\infty x^{z-1} e^{-x} dx = z\Pi(z-1)\end{aligned}$$

- Thus $\Pi(z) = z\Pi(z-1)$, but

$$\Pi(0) = \int_0^\infty e^{-z} dz = [-e^{-x}]_0^\infty$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx = z \int_0^\infty x^{z-1} e^{-x} dx = z\Pi(z-1)\end{aligned}$$

- Thus $\Pi(z) = z\Pi(z-1)$, but

$$\Pi(0) = \int_0^\infty e^{-z} dz = [-e^{-x}]_0^\infty = -e^{-\infty} - (-e^0)$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx = z \int_0^\infty x^{z-1} e^{-x} dx = z\Pi(z-1)\end{aligned}$$

- Thus $\Pi(z) = z\Pi(z-1)$, but

$$\Pi(0) = \int_0^\infty e^{-z} dz = [-e^{-x}]_0^\infty = -e^{-\infty} - (-e^0) = 1$$

Example of Integration by Parts

- Consider

$$\begin{aligned}\Pi(z) &= \int_0^\infty x^z e^{-x} dx = \int_0^\infty x^z \frac{d(-e^{-x})}{dx} dx \\ &= [x^z (-e^{-x})]_0^\infty - \int_0^\infty \frac{dx^z}{dx} (-e^{-x}) dx \\ &= \int_0^\infty (zx^{z-1}) e^{-x} dx = z \int_0^\infty x^{z-1} e^{-x} dx = z\Pi(z-1)\end{aligned}$$

- Thus $\Pi(z) = z\Pi(z-1)$, but

$$\Pi(0) = \int_0^\infty e^{-z} dz = [-e^{-x}]_0^\infty = -e^{-\infty} - (-e^0) = 1$$

- Now

$$\Pi(n) = n\Pi(n-1) = n(n-1)\Pi(n-2) = n(n-1)(n-2)\dots 1 = n!$$

Substitution

- We can make a transformation from x to u

$$\int_a^b f(x) dx = \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i)$$

Substitution

- We can make a transformation from x to $u = u(x)$

$$\int_a^b f(x) dx = \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i)$$

Substitution

- We can make a transformation from x to $u = u(x)$

$$\int_a^b f(x) dx = \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i)$$

Substitution

- We can make a transformation from x to $u = u(x)$

$$\begin{aligned}\int_a^b f(x) dx &= \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i) \\ &= \lim_{\delta u_i \rightarrow 0} \sum_{i=0}^{n-1} f(x(u_i)) \frac{x(u_{i+1}) - x(u_i)}{u_{i+1} - u_i} (u_{i+1} - u_i)\end{aligned}$$

- ★ where u_i is such that $x(u_i) = x_i$ or $u_i = u(x_i)$ where $u(x)$ is the inverse of $x(u)$

Substitution

- We can make a transformation from x to $u = u(x)$

$$\begin{aligned}\int_a^b f(x) dx &= \lim_{\delta x_i \rightarrow 0} \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i) \\ &= \lim_{\delta u_i \rightarrow 0} \sum_{i=0}^{n-1} f(x(u_i)) \frac{x(u_{i+1}) - x(u_i)}{u_{i+1} - u_i} (u_{i+1} - u_i) \\ &= \int_{u(a)}^{u(b)} f(x(u)) \frac{dx(u)}{du} du\end{aligned}$$

★ where u_i is such that $x(u_i) = x_i$ or $u_i = u(x_i)$ where $u(x)$ is the inverse of $x(u)$

★ using $\lim_{\delta u_i \rightarrow 0} \frac{x(u_{i+1}) - x(u_i)}{u_{i+1} - u_i} = \frac{dx(u)}{du}$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$

- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$I(n) = \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du$$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$I(n) = \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du$$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$I(n) = \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du$$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^{\infty} u^{\frac{n-1}{2}} e^{-u} du \end{aligned}$$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^{\infty} u^{\frac{n-1}{2}} e^{-u} du = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2} + 1\right) \end{aligned}$$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^{\infty} u^{\frac{n-1}{2}} e^{-u} du = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2} + 1\right) \end{aligned}$$

- $I(1) = 1$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^{\infty} u^{\frac{n-1}{2}} e^{-u} du = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2} + 1\right) \end{aligned}$$

- $I(1) = 1$, $I(3) = 2 \times 1! = 2$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^{\infty} u^{\frac{n-1}{2}} e^{-u} du = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2} + 1\right) \end{aligned}$$

- $I(1) = 1$, $I(3) = 2 \times 1! = 2$, $I(5) = 2^2 \times 2! = 8$

Example of Integration by Substitution

- We consider $I(n) = \int_0^\infty x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^\infty \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^\infty u^{\frac{n-1}{2}} e^{-u} du = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2} + 1\right) \end{aligned}$$

- $I(1) = 1$, $I(3) = 2 \times 1! = 2$, $I(5) = 2^2 \times 2! = 8$, but
 $I(0) = \Gamma(-1/2)/\sqrt{2}$

Example of Integration by Substitution

- We consider $I(n) = \int_0^{\infty} x^n e^{-x^2/2} dx$
- Let $u(x) = x^2/2$ or $x(u) = \sqrt{2u}$ so that

$$\frac{dx(u)}{du} = \frac{1}{\sqrt{2u}} \quad u(0) = 0 \quad u(\infty) = \infty$$

- Thus

$$\begin{aligned} I(n) &= \int_0^{\infty} \left(\sqrt{2u}\right)^n e^{-u} \frac{1}{\sqrt{2u}} du \\ &= 2^{\frac{n-1}{2}} \int_0^{\infty} u^{\frac{n-1}{2}} e^{-u} du = 2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2} + 1\right) \end{aligned}$$

- $I(1) = 1$, $I(3) = 2 \times 1! = 2$, $I(5) = 2^2 \times 2! = 8$, but
 $I(0) = \Gamma(-1/2)/\sqrt{2}$, $I(2) = \sqrt{2}\Gamma(1/2) = \Gamma(-1/2)/\sqrt{2}$

Changing Variables in Multidimensional Space

- When changing variables in many dimensions $\mathbf{x} \rightarrow \mathbf{u}$ the change of variables involves the Jacobian

$$\int f(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}(\mathbf{u})) |\det(\mathbf{J})| d\mathbf{u}, \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

- E.g. transforming from Cartesian coordinates (x, y) to polar coordinates (r, θ) then $x = r \cos(\theta)$ and $y = r \sin(\theta)$

$$|\det(\mathbf{J})| = \left| \det \begin{pmatrix} \frac{\partial r \cos(\theta)}{\partial r} & \frac{\partial r \cos(\theta)}{\partial \theta} \\ \frac{\partial r \sin(\theta)}{\partial r} & \frac{\partial r \sin(\theta)}{\partial \theta} \end{pmatrix} \right|$$

Changing Variables in Multidimensional Space

- When changing variables in many dimensions $\mathbf{x} \rightarrow \mathbf{u}$ the change of variables involves the Jacobian

$$\int f(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}(\mathbf{u})) |\det(\mathbf{J})| d\mathbf{u}, \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

- E.g. transforming from Cartesian coordinates (x, y) to polar coordinates (r, θ) then $x = r \cos(\theta)$ and $y = r \sin(\theta)$

$$|\det(\mathbf{J})| = \left| \det \begin{pmatrix} \frac{\partial r \cos(\theta)}{\partial r} & \frac{\partial r \cos(\theta)}{\partial \theta} \\ \frac{\partial r \sin(\theta)}{\partial r} & \frac{\partial r \sin(\theta)}{\partial \theta} \end{pmatrix} \right|$$

Changing Variables in Multidimensional Space

- When changing variables in many dimensions $\mathbf{x} \rightarrow \mathbf{u}$ the change of variables involves the Jacobian

$$\int f(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}(\mathbf{u})) |\det(\mathbf{J})| d\mathbf{u}, \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

- E.g. transforming from Cartesian coordinates (x, y) to polar coordinates (r, θ) then $x = r \cos(\theta)$ and $y = r \sin(\theta)$

$$|\det(\mathbf{J})| = \left| \det \begin{pmatrix} \frac{\partial r \cos(\theta)}{\partial r} & \frac{\partial r \cos(\theta)}{\partial \theta} \\ \frac{\partial r \sin(\theta)}{\partial r} & \frac{\partial r \sin(\theta)}{\partial \theta} \end{pmatrix} \right| = \left| \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} \right|$$

Changing Variables in Multidimensional Space

- When changing variables in many dimensions $\mathbf{x} \rightarrow \mathbf{u}$ the change of variables involves the Jacobian

$$\int f(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}(\mathbf{u})) |\det(\mathbf{J})| d\mathbf{u}, \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

- E.g. transforming from Cartesian coordinates (x, y) to polar coordinates (r, θ) then $x = r \cos(\theta)$ and $y = r \sin(\theta)$

$$\begin{aligned} |\det(\mathbf{J})| &= \left| \det \begin{pmatrix} \frac{\partial r \cos(\theta)}{\partial r} & \frac{\partial r \cos(\theta)}{\partial \theta} \\ \frac{\partial r \sin(\theta)}{\partial r} & \frac{\partial r \sin(\theta)}{\partial \theta} \end{pmatrix} \right| = \left| \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} \right| \\ &= r (\cos^2(\theta) + \sin^2(\theta)) \end{aligned}$$

Changing Variables in Multidimensional Space

- When changing variables in many dimensions $\mathbf{x} \rightarrow \mathbf{u}$ the change of variables involves the Jacobian

$$\int f(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}(\mathbf{u})) |\det(\mathbf{J})| d\mathbf{u}, \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

- E.g. transforming from Cartesian coordinates (x, y) to polar coordinates (r, θ) then $x = r \cos(\theta)$ and $y = r \sin(\theta)$

$$\begin{aligned} |\det(\mathbf{J})| &= \left| \det \begin{pmatrix} \frac{\partial r \cos(\theta)}{\partial r} & \frac{\partial r \cos(\theta)}{\partial \theta} \\ \frac{\partial r \sin(\theta)}{\partial r} & \frac{\partial r \sin(\theta)}{\partial \theta} \end{pmatrix} \right| = \left| \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} \right| \\ &= r (\cos^2(\theta) + \sin^2(\theta)) = r \end{aligned}$$

Changing Variables in Multidimensional Space

- When changing variables in many dimensions $\mathbf{x} \rightarrow \mathbf{u}$ the change of variables involves the Jacobian

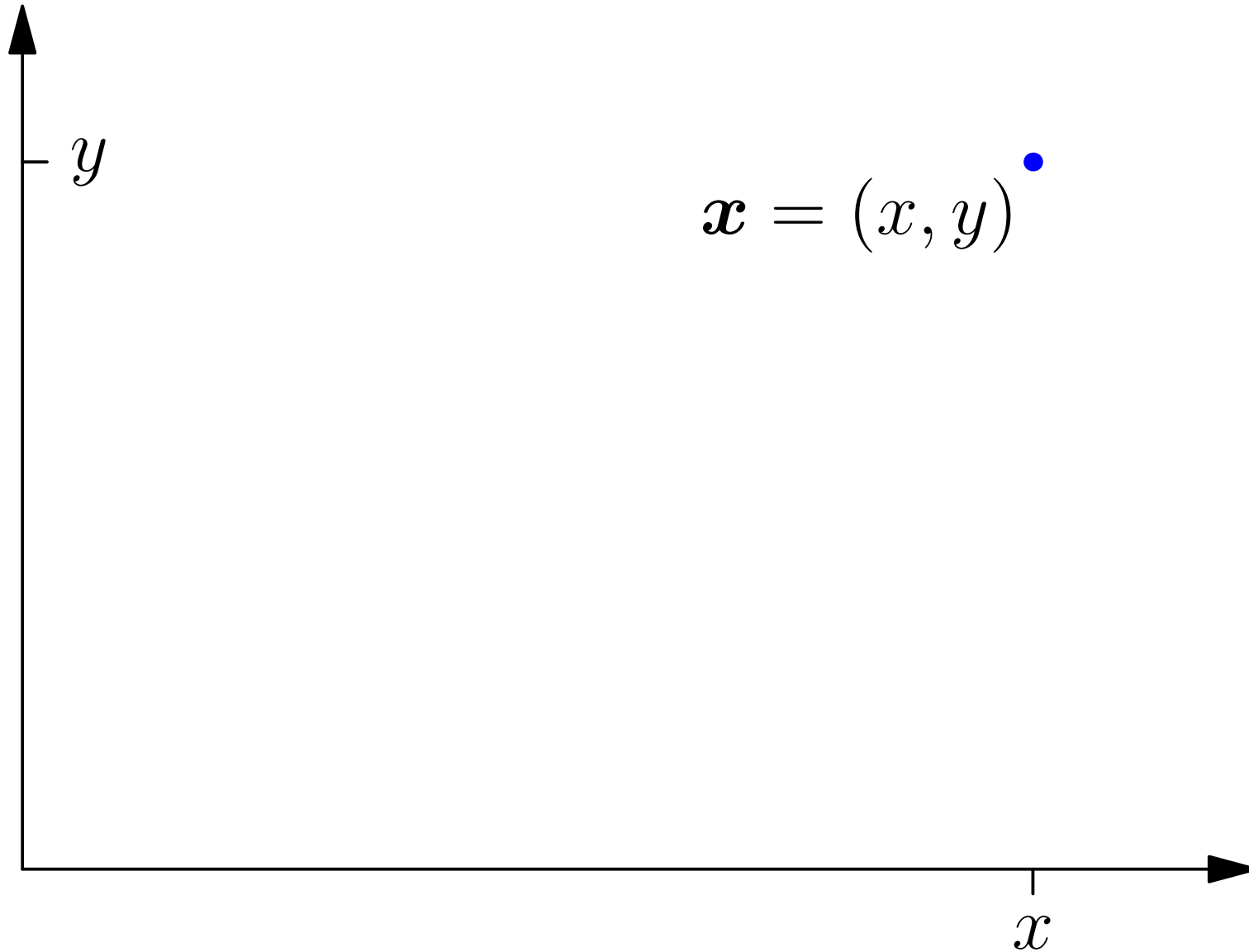
$$\int f(\mathbf{x}) d\mathbf{x} = \int f(\mathbf{x}(\mathbf{u})) |\det(\mathbf{J})| d\mathbf{u}, \quad \mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{pmatrix}$$

- E.g. transforming from Cartesian coordinates (x, y) to polar coordinates (r, θ) then $x = r \cos(\theta)$ and $y = r \sin(\theta)$

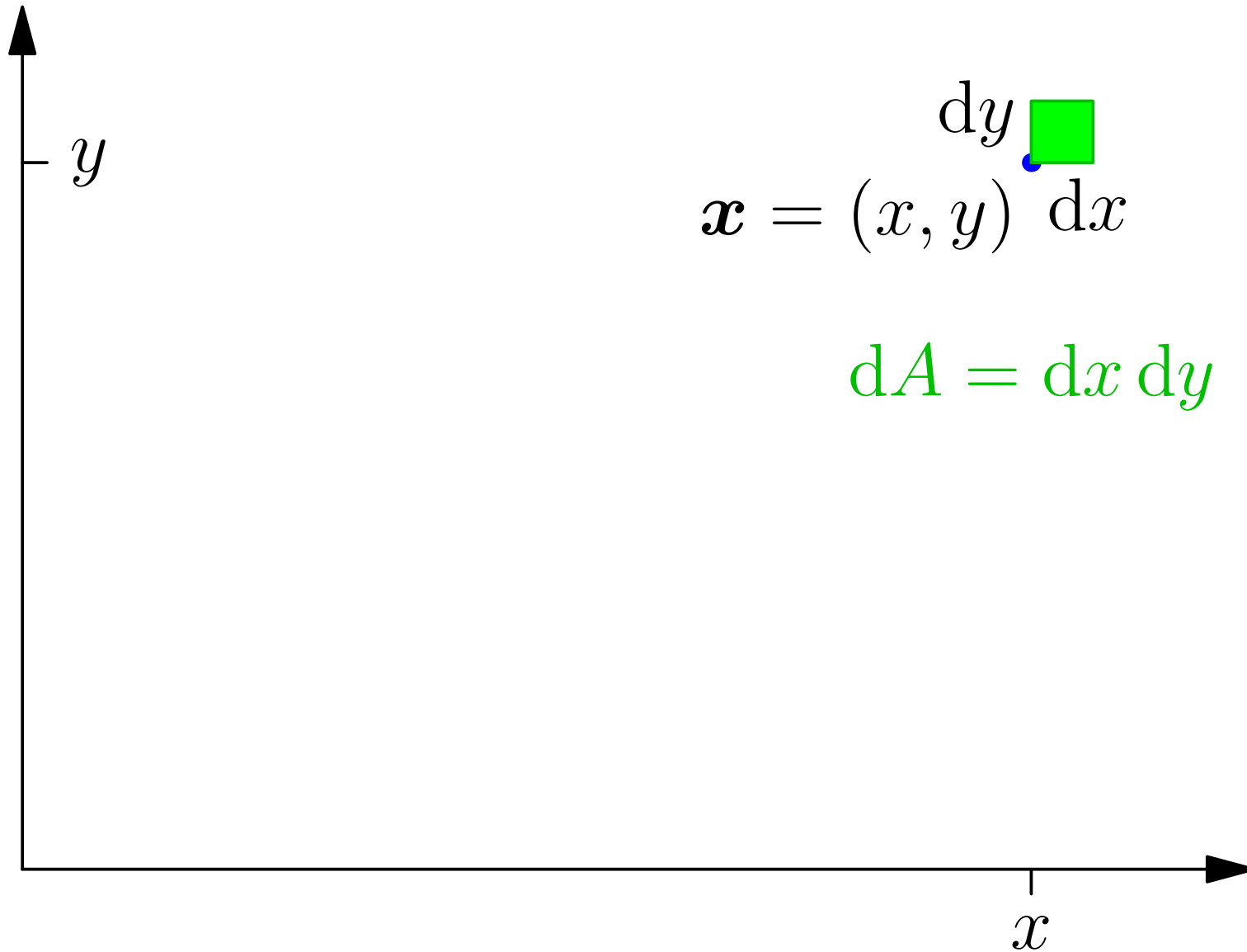
$$\begin{aligned} |\det(\mathbf{J})| &= \left| \det \begin{pmatrix} \frac{\partial r \cos(\theta)}{\partial r} & \frac{\partial r \cos(\theta)}{\partial \theta} \\ \frac{\partial r \sin(\theta)}{\partial r} & \frac{\partial r \sin(\theta)}{\partial \theta} \end{pmatrix} \right| = \left| \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} \right| \\ &= r (\cos^2(\theta) + \sin^2(\theta)) = r \end{aligned}$$

- That is, $dx dy = r dr d\theta$

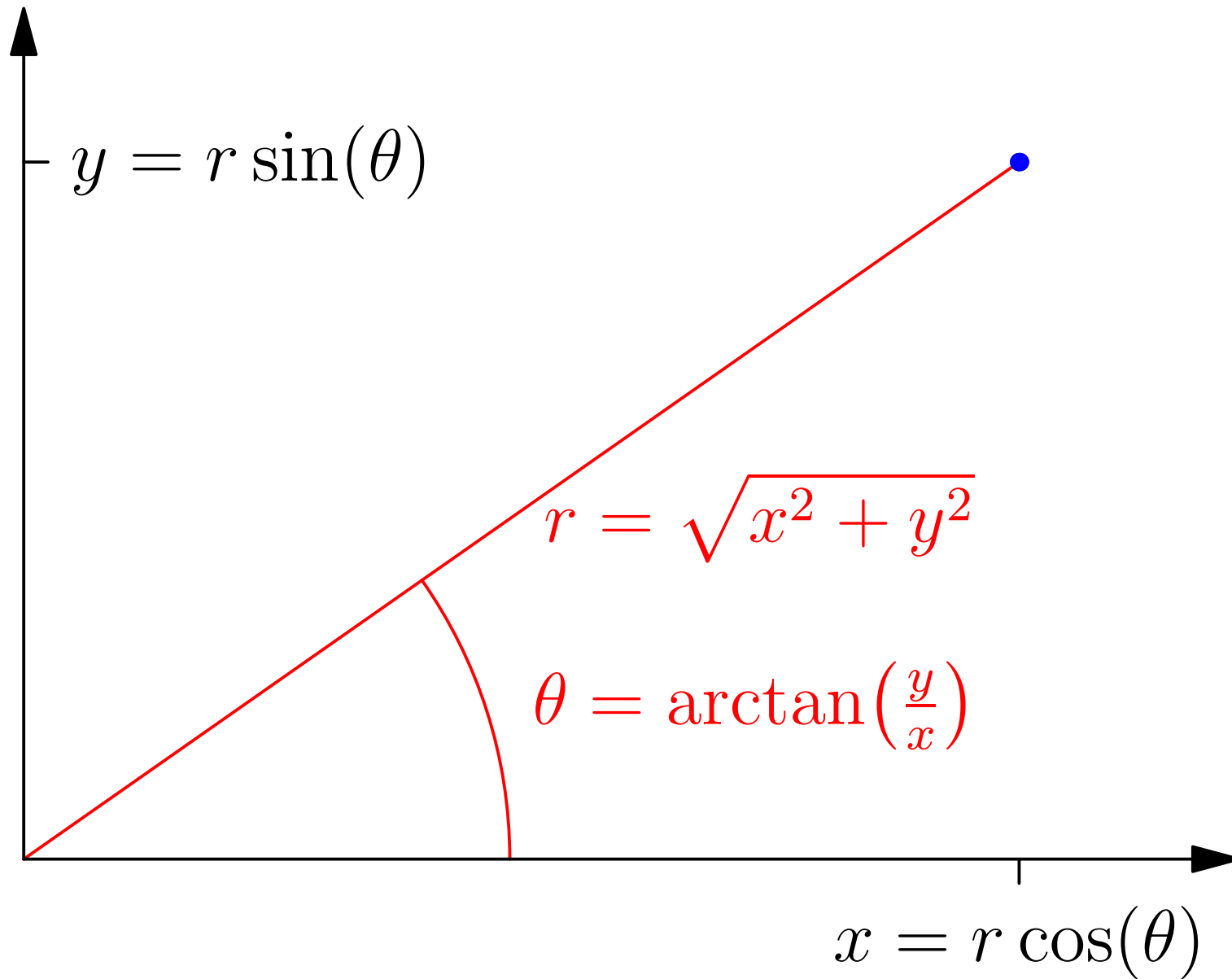
Change of Variables in Pictures



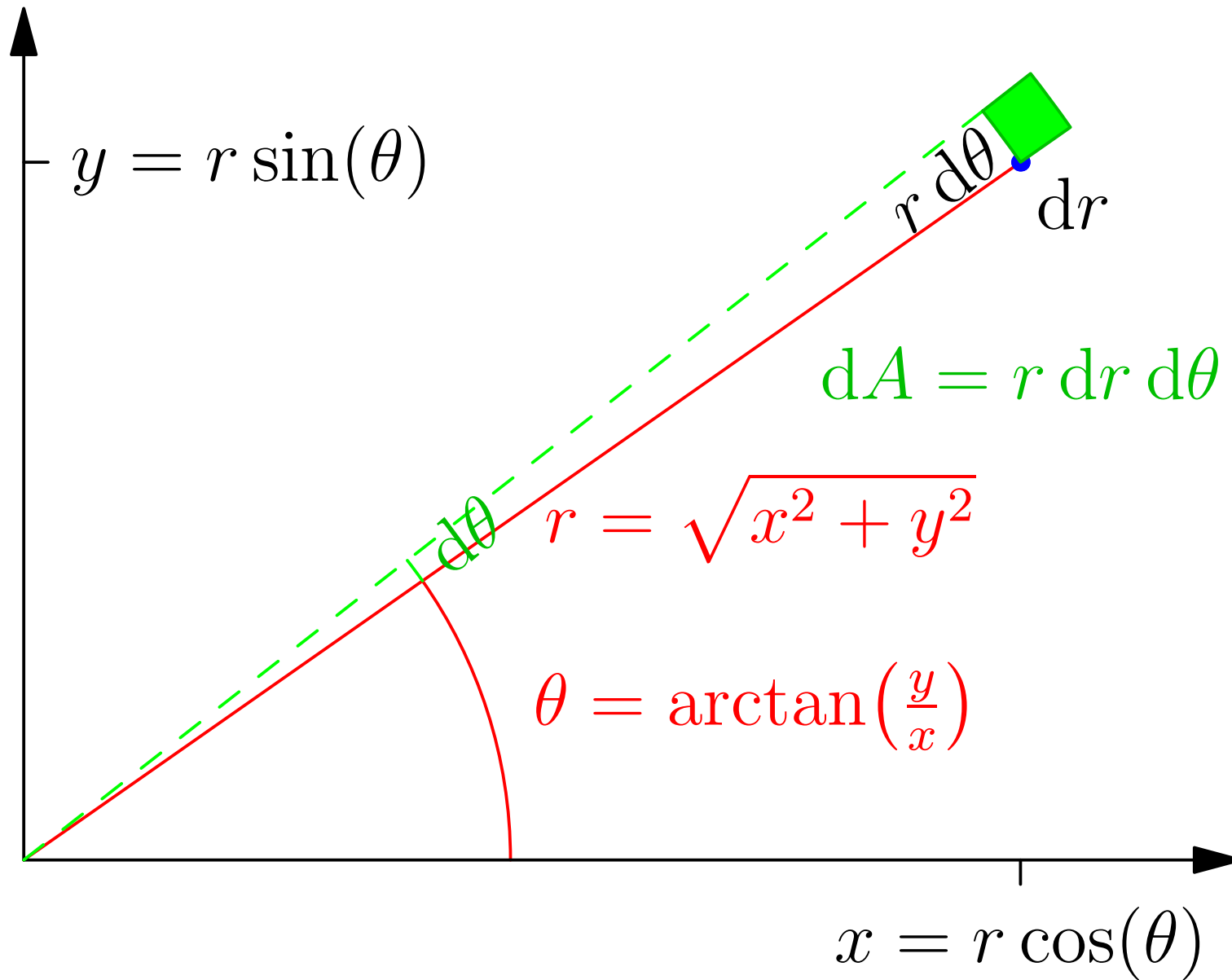
Change of Variables in Pictures



Change of Variables in Pictures



Change of Variables in Pictures



Differentiating Through the Integral

- A trick that sometimes works is differentiating through an integral, e.g. consider finding moments

$$M_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

- We can define a momentum generating function

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

Differentiating Through the Integral

- A trick that sometimes works is differentiating through an integral, e.g. consider finding moments

$$M_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

- We can define a momentum generating function

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

Differentiating Through the Integral

- A trick that sometimes works is differentiating through an integral, e.g. consider finding moments

$$M_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

- We can define a momentum generating function

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

- Then $M_n = Z^{(n)}(0)$

$$\left. \frac{d^n Z(\ell)}{d\ell^n} \right|_{\ell=0} = \int_{-\infty}^{\infty} \left. \frac{d^n e^{\ell x}}{d\ell^n} \right|_{\ell=0} f_X(x) dx$$

Differentiating Through the Integral

- A trick that sometimes works is differentiating through an integral, e.g. consider finding moments

$$M_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

- We can define a momentum generating function

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

- Then $M_n = Z^{(n)}(0)$

$$\left. \frac{d^n Z(\ell)}{d\ell^n} \right|_{\ell=0} = \int_{-\infty}^{\infty} \left. \frac{d^n e^{\ell x}}{d\ell^n} \right|_{\ell=0} f_X(x) dx = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

Differentiating Through the Integral

- A trick that sometimes works is differentiating through an integral, e.g. consider finding moments

$$M_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

- We can define a momentum generating function

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

- Then $M_n = Z^{(n)}(0)$

$$\left. \frac{d^n Z(\ell)}{d\ell^n} \right|_{\ell=0} = \int_{-\infty}^{\infty} \left. \frac{d^n e^{\ell x}}{d\ell^n} \right|_{\ell=0} f_X(x) dx = \int_{-\infty}^{\infty} x^n f_X(x) dx = M_n$$

Cumulant Generating Function

- Note that $e^{\ell x} = 1 + \ell x + \frac{1}{2}\ell^2 x^2 + \frac{1}{3!}\ell^3 x^3 + \dots$

- So

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

- Now using $\log(1 + \epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + \frac{1}{3}\epsilon^3 + \dots$

$$G(\ell) = \log(Z(\ell)) = \ell M_1 + \frac{1}{2}\ell^2 (M_2 - M_1^2) + \frac{1}{3!}\ell^3 (M_3 - 3M_2 M_1 + 2M_1^3) + \dots$$

- So that $\kappa_n = G^{(n)}(0)$, with $\kappa_1 = M_1$ (the mean), $\kappa_2 = M_2 - M_1^2$ (the variance), $\kappa_3 = M_3 - 3M_2 M_1 + 2M_1^3$ (the third cumulant related to the skewness)

Cumulant Generating Function

- Note that $e^{\ell x} = 1 + \ell x + \frac{1}{2}\ell^2 x^2 + \frac{1}{3!}\ell^3 x^3 + \dots$

- So

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx$$

- Now using $\log(1 + \epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + \frac{1}{3}\epsilon^3 + \dots$

$$G(\ell) = \log(Z(\ell)) = \ell M_1 + \frac{1}{2}\ell^2 (M_2 - M_1^2) + \frac{1}{3!}\ell^3 (M_3 - 3M_2 M_1 + 2M_1^3) + \dots$$

- So that $\kappa_n = G^{(n)}(0)$, with $\kappa_1 = M_1$ (the mean), $\kappa_2 = M_2 - M_1^2$ (the variance), $\kappa_3 = M_3 - 3M_2 M_1 + 2M_1^3$ (the third cumulant related to the skewness)

Cumulant Generating Function

- Note that $e^{\ell x} = 1 + \ell x + \frac{1}{2}\ell^2 x^2 + \frac{1}{3!}\ell^3 x^3 + \dots$

- So

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx = 1 + \ell M_1 + \frac{1}{2}\ell^2 M_2 + \frac{1}{3!}\ell^3 M_3 + \dots$$

- Now using $\log(1 + \epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + \frac{1}{3}\epsilon^3 + \dots$

$$G(\ell) = \log(Z(\ell)) = \ell M_1 + \frac{1}{2}\ell^2 (M_2 - M_1^2) + \frac{1}{3!}\ell^3 (M_3 - 3M_2 M_1 + 2M_1^3) + \dots$$

- So that $\kappa_n = G^{(n)}(0)$, with $\kappa_1 = M_1$ (the mean), $\kappa_2 = M_2 - M_1^2$ (the variance), $\kappa_3 = M_3 - 3M_2 M_1 + 2M_1^3$ (the third cumulant related to the skewness)

Cumulant Generating Function

- Note that $e^{\ell x} = 1 + \ell x + \frac{1}{2}\ell^2 x^2 + \frac{1}{3!}\ell^3 x^3 + \dots$

- So

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx = 1 + \ell M_1 + \frac{1}{2}\ell^2 M_2 + \frac{1}{3!}\ell^3 M_3 + \dots$$

- Now using $\log(1 + \epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + \frac{1}{3}\epsilon^3 + \dots$

$$G(\ell) = \log(Z(\ell)) = \ell M_1 + \frac{1}{2}\ell^2 (M_2 - M_1^2) + \frac{1}{3!}\ell^3 (M_3 - 3M_2 M_1 + 2M_1^3) + \dots$$

- So that $\kappa_n = G^{(n)}(0)$, with $\kappa_1 = M_1$ (the mean), $\kappa_2 = M_2 - M_1^2$ (the variance), $\kappa_3 = M_3 - 3M_2 M_1 + 2M_1^3$ (the third cumulant related to the skewness)

Cumulant Generating Function

- Note that $e^{\ell x} = 1 + \ell x + \frac{1}{2}\ell^2 x^2 + \frac{1}{3!}\ell^3 x^3 + \dots$

- So

$$Z(\ell) = \int_{-\infty}^{\infty} e^{\ell x} f_X(x) dx = 1 + \ell M_1 + \frac{1}{2}\ell^2 M_2 + \frac{1}{3!}\ell^3 M_3 + \dots$$

- Now using $\log(1 + \epsilon) = \epsilon - \frac{1}{2}\epsilon^2 + \frac{1}{3}\epsilon^3 + \dots$

$$G(\ell) = \log(Z(\ell)) = \ell M_1 + \frac{1}{2}\ell^2 (M_2 - M_1^2) + \frac{1}{3!}\ell^3 (M_3 - 3M_2 M_1 + 2M_1^3) + \dots$$

- So that $\kappa_n = G^{(n)}(0)$, with $\kappa_1 = M_1$ (the mean), $\kappa_2 = M_2 - M_1^2$ (the variance), $\kappa_3 = M_3 - 3M_2 M_1 + 2M_1^3$ (the third cumulant related to the skewness)

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics
- Interestingly, also there is an algorithm that allows us to integrate a lot of function

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics
- Interestingly, also there is an algorithm that allows us to integrate a lot of function

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics
- Interestingly, also there is an algorithm that allows us to integrate a lot of function

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics (due largely to Cauchy)
- Interestingly, also there is an algorithm that allows us to integrate a lot of function

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics (due largely to Cauchy)—but beyond the scope of this course
- Interestingly, also there is an algorithm that allows us to integrate a lot of function

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics (due largely to Cauchy)—but beyond the scope of this course
- Interestingly, also there is an algorithm that allows us to integrate a lot of function

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics (due largely to Cauchy)—but beyond the scope of this course
- Interestingly, also there is an algorithm that allows us to integrate a lot of function. It is sufficiently complicated that you need to write a computer algorithm of considerable complexity to implement it

More Integration

- Although we have a few tricks, integration is hard
- Surprisingly integration sometimes is easier when carried out in the complex plane
- This is a beautiful part of mathematics (due largely to Cauchy)—but beyond the scope of this course
- Interestingly, also there is an algorithm that allows us to integrate a lot of function. It is sufficiently complicated that you need to write a computer algorithm of considerable complexity to implement it. Most symbolic manipulation packages (e.g. Mathematica) have implemented some part of this algorithm

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z - 1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right
- Although, history has left us with the gamma function instead

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx$$

Special Functions

- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right
- Although, history has left us with the gamma function instead

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx = \Pi(z-1)$$

Special Functions

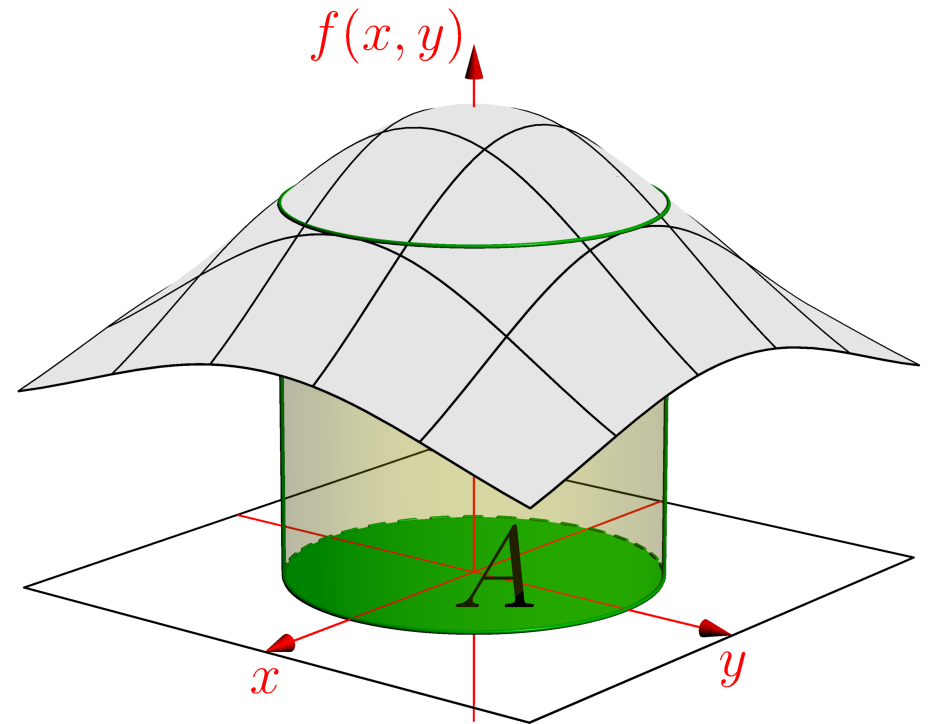
- There are integrals with no known closed form solution
- We saw that $\Pi(z) = \int_0^{\infty} x^z e^{-x} dx$ satisfies $\Pi(z) = z\Pi(z-1)$
- For integer n then $\Pi(n) = n!$, but for general z , the integral $\Pi(z)$ can't be written in terms of elementary functions
- We consider $\Pi(z)$ as a special function in its own right
- Although, history has left us with the gamma function instead

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx = \Pi(z-1)$$

- Other special function defined by integrals exist (e.g. the Bessel , Aire, hypergeometric, elliptic, error functions, . . .)

Outline

1. Defining Integrals
2. Doing Integrals
3. **Gaussian Integrals**



Gaussian Integrals

- Gaussian integrals are integrals involving e^{-x^2} , e.g.

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

$$\int_{-\infty}^{\infty} x^4 e^{-ax^2 - bx} dx$$

- They are important in computing integrals with respect to the normal distribution

$$\mathcal{N}(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

Gaussian Integrals

- Gaussian integrals are integrals involving e^{-x^2} , e.g.

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

$$\int_{-\infty}^{\infty} x^4 e^{-ax^2 - bx} dx$$

- They are important in computing integrals with respect to the normal distribution

$$\mathcal{N}(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

Gaussian Integrals

- Gaussian integrals are integrals involving e^{-x^2} , e.g.

$$\int_{-\infty}^{\infty} e^{-x^2} dx \qquad \int_{-\infty}^{\infty} x^4 e^{-ax^2 - bx} dx$$

- They are important in computing integrals with respect to the normal distribution

$$\mathcal{N}(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

- The great news is that these integrals are all doable

Gaussian Integrals

- Gaussian integrals are integrals involving e^{-x^2} , e.g.

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

$$\int_{-\infty}^{\infty} x^4 e^{-ax^2 - bx} dx$$

- They are important in computing integrals with respect to the normal distribution

$$\mathcal{N}(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

- The great news is that these integrals are all doable
- The bad news is that they are quite tricky to do

The Gaussian Integral

- The integral over a Gaussian is surprisingly difficult

$$I_1 = \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

- There is a nice trick which is to consider

$$I_1^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy$$

- Making the change of variables $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$ (so that $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $x^2 + y^2 = r^2$)

$$I_1^2 = \int_0^{2\pi} d\theta \int_0^{\infty} r e^{-r^2/2} dr = 2\pi \int_0^{\infty} r e^{-r^2/2} dr$$

The Gaussian Integral

- The integral over a Gaussian is surprisingly difficult

$$I_1 = \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

- There is a nice trick which is to consider

$$I_1^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy$$

- Making the change of variables $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$ (so that $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $x^2 + y^2 = r^2$)

$$I_1^2 = \int_0^{2\pi} d\theta \int_0^{\infty} r e^{-r^2/2} dr = 2\pi \int_0^{\infty} r e^{-r^2/2} dr$$

The Gaussian Integral

- The integral over a Gaussian is surprisingly difficult

$$I_1 = \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

- There is a nice trick which is to consider

$$I_1^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy$$

- Making the change of variables $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$ (so that $x = r \cos(\theta)$, $y = r \sin(\theta)$ and $x^2 + y^2 = r^2$)

$$I_1^2 = \int_0^{2\pi} d\theta \int_0^{\infty} r e^{-r^2/2} dr = 2\pi \int_0^{\infty} r e^{-r^2/2} dr$$

The Gaussian Integral Continued

- From before

$$I_1^2 = 2\pi \int_0^\infty r e^{-r^2/2} dr$$

- Finally let $u = r^2/2$ so that $du/dr = r$ or $du = r dr$ we get

$$I_1^2 = 2\pi \int_0^\infty e^{-u} du = 2\pi$$

The Gaussian Integral Continued

- From before

$$I_1^2 = 2\pi \int_0^\infty r e^{-r^2/2} dr$$

- Finally let $u = r^2/2$ so that $du/dr = r$ or $du = r dr$ we get

$$I_1^2 = 2\pi \int_0^\infty e^{-u} du = 2\pi$$

- So that $I_1 = \sqrt{2\pi}$

The Gaussian Integral Continued

- From before

$$I_1^2 = 2\pi \int_0^\infty r e^{-r^2/2} dr$$

- Finally let $u = r^2/2$ so that $du/dr = r$ or $du = r dr$ we get

$$I_1^2 = 2\pi \int_0^\infty e^{-u} du = 2\pi$$

- So that $I_1 = \sqrt{2\pi}$
- Incidentally, $I_1 = \sqrt{2}\Pi(-1/2)$ so $\Pi(-1/2) = \Gamma(1/2) = \sqrt{\pi}$

Normal Distribution

- We consider

$$I_2 = \int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)} dx$$

- Making the change of variables $z = (x - \mu)/\sigma$ so that $dz = dx/\sigma$ or $dx = \sigma dz$. Then

$$I_2 = \sigma \int_{-\infty}^{\infty} e^{-z^2/2} dz = \sigma I_1 = \sqrt{2\pi} \sigma$$

- Note that the *probability density function* (PDF) for a normally distributed random variable is given by

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

Normal Distribution

- We consider

$$I_2 = \int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)} dx$$

- Making the change of variables $z = (x - \mu)/\sigma$ so that $dz = dx/\sigma$ or $dx = \sigma dz$. Then

$$I_2 = \sigma \int_{-\infty}^{\infty} e^{-z^2/2} dz = \sigma I_1 = \sqrt{2\pi} \sigma$$

- Note that the *probability density function* (PDF) for a normally distributed random variable is given by

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

Normal Distribution

- We consider

$$I_2 = \int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)} dx$$

- Making the change of variables $z = (x - \mu)/\sigma$ so that $dz = dx/\sigma$ or $dx = \sigma dz$. Then

$$I_2 = \sigma \int_{-\infty}^{\infty} e^{-z^2/2} dz = \sigma I_1 = \sqrt{2\pi} \sigma$$

- Note that the *probability density function* (PDF) for a normally distributed random variable is given by

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

Multi-dimensional Gaussians

- Consider

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\|\mathbf{x}\|_2^2} dx_1 \cdots dx_n$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$

- Note that $\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2$ and using $e^{\sum_i a_i} = \prod_i e^{a_i}$

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2} dx_1 \cdots dx_n$$

Multi-dimensional Gaussians

- Consider

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\|\mathbf{x}\|_2^2} dx_1 \cdots dx_n$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$

- Note that $\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2$ and using $e^{\sum_i a_i} = \prod_i e^{a_i}$

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2} dx_1 \cdots dx_n$$

Multi-dimensional Gaussians

- Consider

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\|\mathbf{x}\|_2^2} dx_1 \cdots dx_n$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$

- Note that $\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2$ and using $e^{\sum_i a_i} = \prod_i e^{a_i}$

$$\begin{aligned} I_3 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2} dx_1 \cdots dx_n \\ &= \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-x_i^2/2} dx_i \end{aligned}$$

Multi-dimensional Gaussians

- Consider

$$I_3 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\|\mathbf{x}\|_2^2} dx_1 \cdots dx_n$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$

- Note that $\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2$ and using $e^{\sum_i a_i} = \prod_i e^{a_i}$

$$\begin{aligned} I_3 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \sum_{i=1}^n x_i^2} dx_1 \cdots dx_n \\ &= \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-x_i^2/2} dx_i = \prod_{i=1}^n \sqrt{2\pi} = (2\pi)^{n/2} \end{aligned}$$

Full Multi-variate Normal

- Consider

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Xi}^{-1}(\mathbf{x}-\boldsymbol{\mu})} d\mathbf{x}_1 \cdots d\mathbf{x}_n$$

- Let $\boldsymbol{\Xi}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$ and make the change of variables $\mathbf{y} = \mathbf{V}^T(\mathbf{x} - \boldsymbol{\mu})$
- The Jacobian \mathbf{J} has elements (note that $\mathbf{x} = \mathbf{V}\mathbf{y} + \boldsymbol{\mu}$)

$$J_{ij} = \frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j} \left(\sum_{k=1}^n V_{ik} y_k + \mu_i \right) = V_{ij}$$

- So that $\mathbf{J} = \mathbf{V}$ and consequently $|\det(\mathbf{J})| = |\det(\mathbf{V})| = 1$ then

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Lambda}^{-1} \mathbf{y}} dy_1 \cdots dy_n = \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-y_i^2 / (2\lambda_i)} dy_i = \prod_i \sqrt{2\pi\lambda_i}$$

Full Multi-variate Normal

- Consider

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Xi}^{-1}(\mathbf{x}-\boldsymbol{\mu})} dx_1 \cdots dx_n$$

- Let $\boldsymbol{\Xi}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$ and make the change of variables $\mathbf{y} = \mathbf{V}^T(\mathbf{x} - \boldsymbol{\mu})$
- The Jacobian \mathbf{J} has elements (note that $\mathbf{x} = \mathbf{V}\mathbf{y} + \boldsymbol{\mu}$)

$$J_{ij} = \frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j} \left(\sum_{k=1}^n V_{ik} y_k + \mu_i \right) = V_{ij}$$

- So that $\mathbf{J} = \mathbf{V}$ and consequently $|\det(\mathbf{J})| = |\det(\mathbf{V})| = 1$ then

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Lambda}^{-1}\mathbf{y}} dy_1 \cdots dy_n = \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-y_i^2/(2\lambda_i)} dy_i = \prod_i \sqrt{2\pi\lambda_i}$$

Full Multi-variate Normal

- Consider

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Xi}^{-1}(\mathbf{x}-\boldsymbol{\mu})} dx_1 \cdots dx_n$$

- Let $\boldsymbol{\Xi}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$ and make the change of variables $\mathbf{y} = \mathbf{V}^T(\mathbf{x} - \boldsymbol{\mu})$
- The Jacobian \mathbf{J} has elements (note that $\mathbf{x} = \mathbf{V}\mathbf{y} + \boldsymbol{\mu}$)

$$J_{ij} = \frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j} \left(\sum_{k=1}^n V_{ik} y_k + \mu_i \right) = V_{ij}$$

- So that $\mathbf{J} = \mathbf{V}$ and consequently $|\det(\mathbf{J})| = |\det(\mathbf{V})| = 1$ then

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Lambda}^{-1} \mathbf{y}} dy_1 \cdots dy_n = \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-y_i^2/(2\lambda_i)} dy_i = \prod_i \sqrt{2\pi\lambda_i}$$

Full Multi-variate Normal

- Consider

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Xi}^{-1}(\mathbf{x}-\boldsymbol{\mu})} dx_1 \cdots dx_n$$

- Let $\boldsymbol{\Xi}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$ and make the change of variables $\mathbf{y} = \mathbf{V}^T(\mathbf{x} - \boldsymbol{\mu})$
- The Jacobian \mathbf{J} has elements (note that $\mathbf{x} = \mathbf{V}\mathbf{y} + \boldsymbol{\mu}$)

$$J_{ij} = \frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j} \left(\sum_{k=1}^n V_{ik} y_k + \mu_i \right) = V_{ij}$$

- So that $\mathbf{J} = \mathbf{V}$ and consequently $|\det(\mathbf{J})| = |\det(\mathbf{V})| = 1$ then

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Lambda}^{-1}\mathbf{y}} dy_1 \cdots dy_n = \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-y_i^2/(2\lambda_i)} dy_i = \prod_i \sqrt{2\pi\lambda_i}$$

Full Multi-variate Normal

- Consider

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Xi}^{-1}(\mathbf{x}-\boldsymbol{\mu})} dx_1 \cdots dx_n$$

- Let $\boldsymbol{\Xi}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$ and make the change of variables $\mathbf{y} = \mathbf{V}^T(\mathbf{x} - \boldsymbol{\mu})$
- The Jacobian \mathbf{J} has elements (note that $\mathbf{x} = \mathbf{V}\mathbf{y} + \boldsymbol{\mu}$)

$$J_{ij} = \frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j} \left(\sum_{k=1}^n V_{ik} y_k + \mu_i \right) = V_{ij}$$

- So that $\mathbf{J} = \mathbf{V}$ and consequently $|\det(\mathbf{J})| = |\det(\mathbf{V})| = 1$ then

$$I_4 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Lambda}^{-1}\mathbf{y}} dy_1 \cdots dy_n = \prod_{i=1}^n \int_{-\infty}^{\infty} e^{-y_i^2/(2\lambda_i)} dy_i = \prod_i \sqrt{2\pi\lambda_i}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top)$$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V}^\top)$$

using $\det(\mathbf{A}\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V}^\top) = \det(\mathbf{\Lambda})$$

using $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{V}) = 1$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V}^\top) = \det(\mathbf{\Lambda}) = \prod_{i=1}^n \lambda_i$$

using $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{V}) = 1$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V}^\top) = \det(\mathbf{\Lambda}) = \prod_{i=1}^n \lambda_i$$

using $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{V}) = 1$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i} = (2\pi)^{n/2} \sqrt{\det(\Xi)}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V}^\top) = \det(\mathbf{\Lambda}) = \prod_{i=1}^n \lambda_i$$

using $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{V}) = 1$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i} = (2\pi)^{n/2} \sqrt{\det(\Xi)}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Determinants

- Using the facts, that $\Xi = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ then

$$\det(\Xi) = \det(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top) = \det(\mathbf{V})\det(\mathbf{\Lambda})\det(\mathbf{V}^\top) = \det(\mathbf{\Lambda}) = \prod_{i=1}^n \lambda_i$$

using $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{V}) = 1$

- Recall $I_4 = \prod_i \sqrt{2\pi\lambda_i} = (2\pi)^{n/2} \sqrt{\det(\Xi)}$
- We note for an $n \times n$ matrix \mathbf{M} then $\det(c\mathbf{M}) = c^n \det(\mathbf{M})$ so that

$$I_4 = (2\pi)^{n/2} \sqrt{\det(\Xi)} = \sqrt{\det(2\pi\Xi)}$$

- Finally, we get that for the PDF of a normal to integrate to 1

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Xi) = \frac{1}{\sqrt{\det(2\pi\Xi)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Xi^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard
- Normal distributions lucky almost always give raise to integrals that can be computed in closed form

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard
- Normal distributions lucky almost always give raise to integrals that can be computed in closed form

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard
- Normal distributions lucky almost always give raise to integrals that can be computed in closed form

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard—often impossible
- Normal distributions lucky almost always give raise to integrals that can be computed in closed form

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard—often impossible
- Normal distributions lucky almost always give rise to integrals that can be computed in closed form

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard—often impossible
- Normal distributions lucky almost always give rise to integrals that can be computed in closed form, **although often it requires quite a bit of work**

Summary

- Integration is extra-ordinarily useful as a tool of analysis
- It occurs when you work with probabilities densities for continuous random variables
- Integration is beautiful, but hard—often impossible
- Normal distributions lucky almost always give rise to integrals that can be computed in closed form, although often it requires quite a bit of work
- Making friends with integration will give you a super-power that not too many people share