8 Calculus III

8.1 Optimization

We often want to know where our functions are at their maximum or minimum. We do this in two steps. First, wherever the first derivative is at 0, means we are either at a minimum or maximum. Second, depending upon the direction of the second derivative, we can tell if we are at a maximum or minimum.

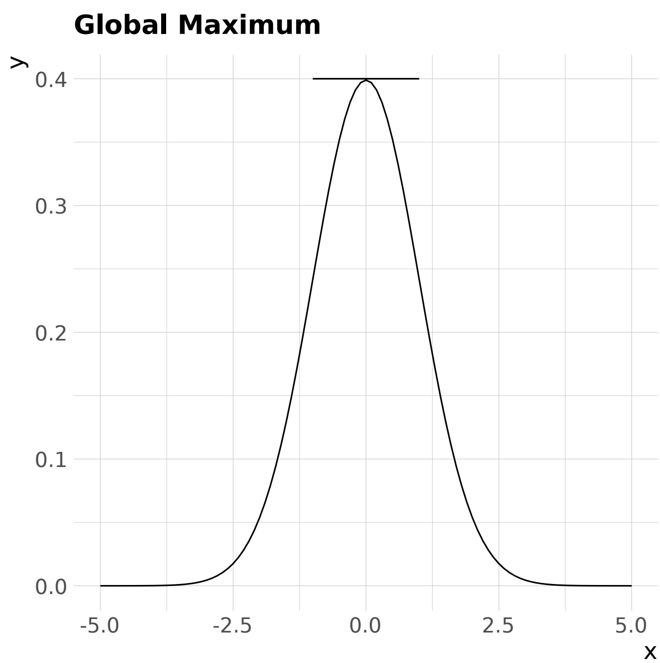


Figure 2: Visualizing how a derivative is equal to 0 at a function's maximum or minimum.

- **First-order condition**: $f'(x) = 0$
  - Maximum or minimum
  - This assumes that $f(x)$ is differentiable at all values – it is a continuous function.
- **Second-order condition**: $f''(x)$, where > 0 means minimum, < 0 means maximum, and $= 0$ an inflection point.

$$\begin{aligned} f(x) &= x^2 - 4x - 1 \\ \text{FOC: } f'(x) &= 2x - 4 \\ x &= 2 \\ \text{SOC: } f''(x) &= 2 \Rightarrow \text{minimum} \end{aligned}$$

What if we have multiple variables?

- Let $f(x) = -x_1^2 + x_1x_2 - x_2^2$
  - FOC:
    - $f_{x_1} = -2x_1 + x_2 = 0$
    - $f_{x_2} = x_1 - 2x_2 = 0$

    $$\left[\begin{array}{cc|c} -2 & 1 & 0 \\ 1 & -2 & 0 \end{array}\right]$$

      - It turns out $x_1 = 0$, $x_2 = 0$
  - SOC: We need to build the hessian:

  $$\begin{bmatrix} \frac{d^2 f}{dx_1 x_1} & \frac{d^2}{dx_1 x_2} \\ \\ \frac{d^2 f}{dx_2 x_1} & \frac{d^2 f}{dx_2 x_2} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

- Now, we ask whether these points are minima, maxima, indeterminate, or saddle points.
  - We calculate the determinants of each "principal minor".
  - $\text{PM}_1 = -2$, $D_1 = -2$
  - $\text{PM}_2 = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$, $D_2 = (-2 \cdot -2) - (1 \cdot 1) = 3$
- The hessian is:
  - Positive definite: $D_1 > 0$, $D_2 > 0$, strictly local minima
  - Negative definite: $D_1 < 0$, $D_2 > 0$, strictly local maxima
    - Start negative and then alternate signs
  - Positive semi-definite: $D_1 \geqslant 0$, $D_2 \geqslant 0$
  - Negative semi-definite: $D_1 \leqslant 0$, $D_2 \geqslant 0$
- For larger matrices?
  - $\begin{bmatrix} . \end{bmatrix}$, $\text{PM}_1 =$ determinant of 1×1
  - $\begin{bmatrix} . & . \\ . & . \end{bmatrix}$, $\text{PM}_2 =$ determinant of 2×2
  - $\begin{bmatrix} . & . & . \\ . & . & . \\ . & . & . \end{bmatrix}$, $\text{PM}_3 =$ determinant of 3×3

- If we have a 3×3 matrix:
  - PD: $D_1 > 0$, $D_2 > 0$, $D_3 > 0$
  - ND: $D_1 < 0$, $D_2 > 0$, $D_3 < 0$
  - PSD: $D_1 \geqslant 0$, $D_2 \geqslant 0$, $D_3 \geqslant 0$
  - NSD: $D_1 \leqslant 0$, $D_2 \geqslant 0$, $D_3 \leqslant 0$

8.2 Constrained Optimization – Lagrange Multiplier

Sometimes, we have to find the maxima or minima under set conditions.

$$\begin{aligned} &\text{Max } f(x_1, x_2) \text{ s.t. } g(x_1, x_2) = 0 \\ &\text{Max } f(x_1, x_2) - \lambda g(x_1, x_2) \end{aligned}$$

- λ is us creating a new variable.
- FOC:

$$\begin{aligned} f_{x_1} - \lambda g_{x_1} &= 0 \\ f_{x_2} - \lambda g_{x_1} &= 0 \\ g(x_1, x_2) &= 0 \end{aligned}$$

- Then:

$$\begin{aligned} f(x_1, x_2) =& 36 - x_1^2 - x_2^2,\ g(x_1, x_2) = x_1 + 7x_2 - 25 \\ f(x_1, x_2, \lambda) =& 36 - x_1^2 - x_2^2 - \lambda(x1 + 7x_2 - 25) \\ f_{x_1} &= -2x_1 - \lambda = 0, \Rightarrow x_1 = -\lambda/2 \\ f_{x_2} &= -2x^2 - 7\lambda = 0, \Rightarrow x_2 = -7\lambda/2 \\ g(x_1, x_2) &= -\lambda/2 + -7(7\lambda/2) - 25 = 0 \\ \lambda &= -1 \\ \text{Therefore: } x_1 &= \frac{1}{2},\ x_2 = \frac{7}{2} \end{aligned}$$

8.3 Integration

- $f(x) = 40x$
- $F(x) = 20x^2 + c$

  - This is the 'antiderivative', $F'(x) = f(x)$, because if $F(X)$ is derived then it returns $f(x)$.
  - We add c because any constant will turn to 0 when the derivative is taken.

- Powerfully, antidifferentiation lets us solve for the value of y, given dy/dx.

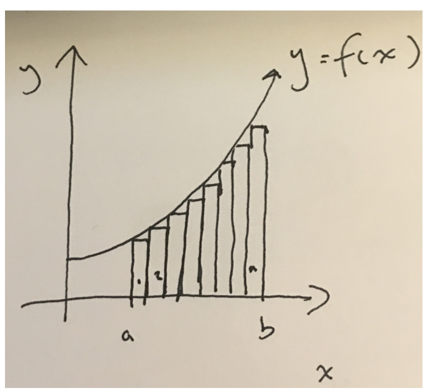
- **Rules of integration**

  - Power rule: $f(x) = x^n$, $F(x) = \frac{1}{n+1}x^{n+1} + c$
  - Exponent rule: $f(x) = e^x$, $F(x) = e^x + c$
    - But: $f(x) = e^a x$, $F(x) = \frac{1}{a}e^{ax} + c$
  - Logarithm rule: $f(x) = \frac{1}{x}$, $F(x) = ln(x) + c$
  - Chain rule: $f(x) = g'(x)e^{g(x)}$, $F(x) = e^{g(x)} + c$
  - Chain + log: $f(x) = \frac{g'(x)}{g(x)}$, $F(x) = ln|g(x)| + c$
  - Sum rule: $f(x) = g(x) + h(x)$, $F(x) = G(x) + H(x)$
  - Constant rule: $f(x) = kg(x)$, $F(x) = kG(x)$
  - Integral of a constant: $f(x) = k$, $F(x) = kx + c$

We integrate so that we can add continuously. The integral is the limit of the sum, where we add all slices under the curve until we reach infinity. We end up summing the values of infinitesimally small ranges. This is the idea behind Riemann sums.

8.4 Areas and Riemann Sums

Consider the following function



where $\sum_{u=1}^{n} f(x_i)\Delta x$ and $\Delta x = \frac{b-a}{n}$. This equals the sum of the area of all the rectangles under $f(x)$. The closer n gets to ∞, then the closer the sum of all the areas gets to the area under $f(x)$ from a to b.

It also turns out that

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x = \int_b^a f(x)dx$$