Mathematical Programming

CSCI 545 Introduction to Robotics
Instructor: Stefanos Nikolaidis

Resources : *16-811: Math Fundamentals for Robotics by Michael Erdmann, Carnegie Mellon University*

Optimization Problems

- How can a robot reach a specific object while avoiding obstacles?
- How can a robot go to its destination while minimizing energy consumption?

Basic Problem

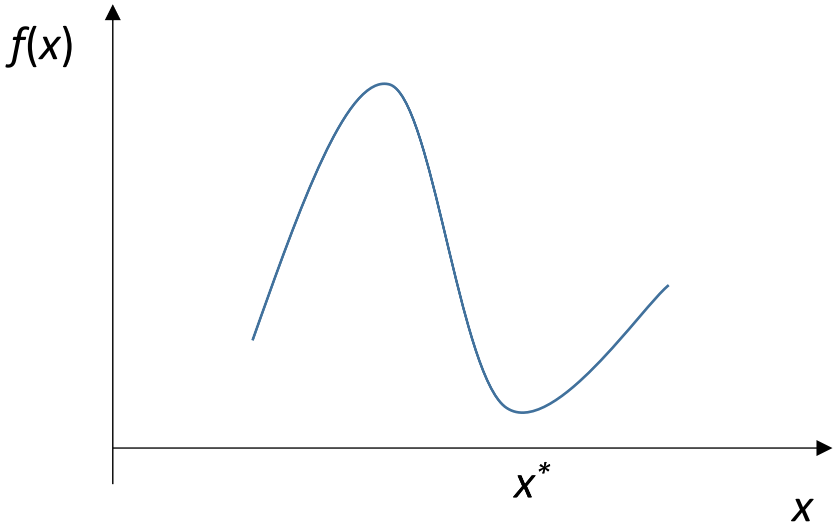
- Given $f : \mathbb{R}^n \rightarrow \mathbb{R}$ find $\min f$

Basic Problem

- Given $f : \mathbb{R}^n \rightarrow \mathbb{R}$ find min f

- Let's start with $n = 1$

$$f : \mathbb{R} \rightarrow \mathbb{R}$$



Basic Problem

- Find min $f : \mathbb{R} \to \mathbb{R}$

- We compute the critical set of *f*:

$$C_f = \{x | f'(x) = 0\}$$

- We identify minima by finding where $f''(x) \geq 0$

Basic Problem

- What about the n dimensional case?

Basic Problem

- What about the n dimensional case?
- Sufficient conditions for a relative (local) case:
- We let $f : \mathbb{R}^n \to \mathbb{R}$ be C^2
- We let x^* be a relative minimum of f :
  - Then: i) $\nabla f(x^*) = 0$

    ii) $\nabla^2 f(x^*) \succ 0$ (positive definite)

Basic Problem

$$\nabla^2 f(x) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_n} \\[2.2ex] \dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_n} \\[2.2ex] \vdots & \vdots & \ddots & \vdots \\[2.2ex] \dfrac{\partial^2 f}{\partial x_n\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix},$$

Example

$$c \in \mathbb{R}$$

$$b \in \mathbb{R}^n$$

$$A = (a_{ij}) \in \mathbb{R}^{n \times n}, a_{ij} = a_{ji}, A \succ 0$$

$$f(x) = c + b^T x + 0.5 x^T A x$$

Example

$$c \in \mathbb{R}$$

$$b \in \mathbb{R}^n$$

$$A = (a_{ij}) \in \mathbb{R}^{n \times n}, a_{ij} = a_{ji}, A \succ 0$$

$$f(x) = c + b^T x + 0.5 x^T A x$$

$$\nabla f(x)?$$

Example

$$c \in \mathbb{R}$$

$$b \in \mathbb{R}^n$$

$$A = (a_{ij}) \in \mathbb{R}^{n \times n}, a_{ij} = a_{ji}, A \succ 0$$

$$f(x) = c + b^T x + 0.5 x^T A x$$

$$\nabla f(x) = b + Ax$$

Example (n=1)

$$a > 0$$

$$f(x) = c + bx + \frac{1}{2}ax^2$$

Example (n=1)

$$a > 0$$

$$f(x) = c + bx + \frac{1}{2}ax^2$$

$$f'(x) = b + ax$$

Example (n=1)

$$a > 0$$

$$f(x) = c + bx + \frac{1}{2}ax^2$$

$$f'(x) = b + ax$$

$$f''(x) = a > 0$$

Example (n=1)

$$a > 0$$

$$f(x) = c + bx + \frac{1}{2}ax^2$$

$$f'(x) = b + ax$$

$$f''(x) = a > 0$$

$$f'(x^*) = 0$$

$$x^* = -\frac{b}{a}$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x) = c + b_1 x_1 + b_2 x_2 + \frac{1}{2} \begin{bmatrix} x_1 a_{11} + x_2 a_{21} \\ x_1 a_{12} + x_2 a_{22} \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + x_1 x_2 a_{21} + x_1 x_2 a_{12} + x_2^2 a_{22})$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x) = c + b_1 x_1 + b_2 x_2 + \frac{1}{2} \begin{bmatrix} x_1 a_{11} + x_2 a_{21} \\ x_1 a_{12} + x_2 a_{22} \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + x_1 x_2 a_{21} + x_1 x_2 a_{12} + x_2^2 a_{22})$$

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + 2 x_1 x_2 a_{12} + x_2^2 a_{22})$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + 2x_1 x_2 a_{12} + x_2^2 a_{22})$$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + 2x_1 x_2 a_{12} + x_2^2 a_{22})$$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 + x_1 a_{11} + x_2 a_{12} \\ b_2 + x_2 a_{12} + x_1 a_{21} \end{bmatrix}$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + 2x_1 x_2 a_{12} + x_2^2 a_{22})$$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 + x_1 a_{11} + x_2 a_{12} \\ b_2 + x_2 a_{12} + x_1 a_{21} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Example (n=2)

- Say $x = (x_1, x_2)$:

$$f(x) = c + b_1 x_1 + b_2 x_2 + 0.5(x_1^2 a_{11} + 2x_1 x_2 a_{12} + x_2^2 a_{22})$$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 + x_1 a_{11} + x_2 a_{12} \\ b_2 + x_2 a_{12} + x_1 a_{21} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= b + Ax$$

Example (n=2)

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 + x_1 a_{11} + x_2 a_{12} \\ b_2 + x_2 a_{12} + x_1 a_{21} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= b + Ax$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

Example (n=2)

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 + x_1 a_{11} + x_2 a_{12} \\ b_2 + x_2 a_{12} + x_1 a_{21} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= b + Ax$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} = A$$

Example (n=2)

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} b_1 + x_1 a_{11} + x_2 a_{12} \\ b_2 + x_2 a_{12} + x_1 a_{21} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= b + Ax$$

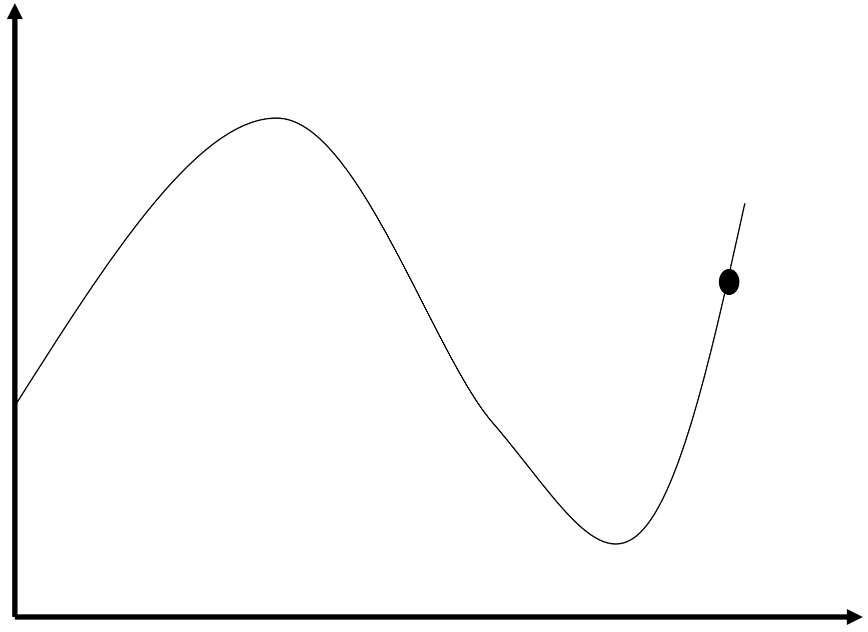
$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} = A$$

There is a local minimum that is the solution to the equation: $Ax^* = -b$

Since it is the only one, this is a global minimum.

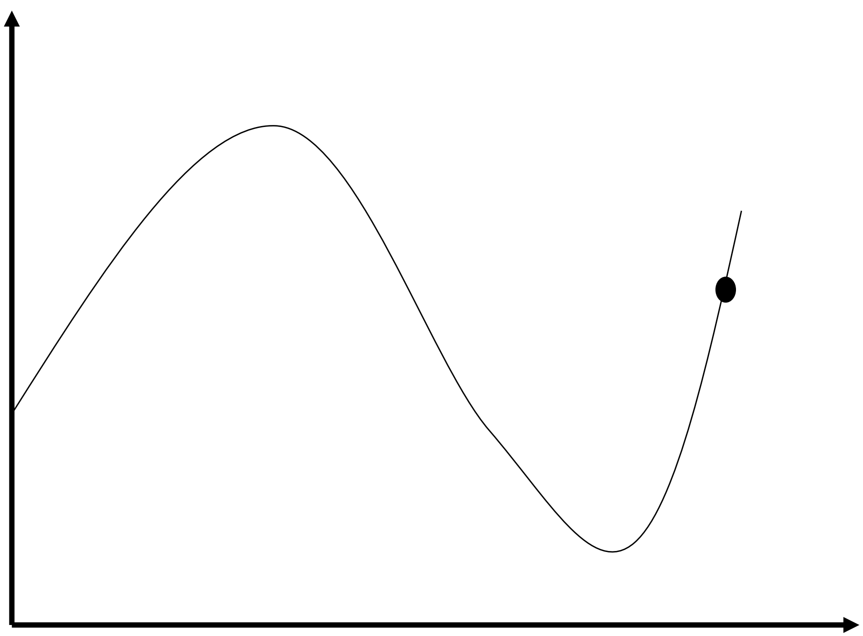
Gradient Descent

- Given $f : \mathbb{R} \to \mathbb{R}$



Gradient Descent

- Given $f : \mathbb{R} \to \mathbb{R}$

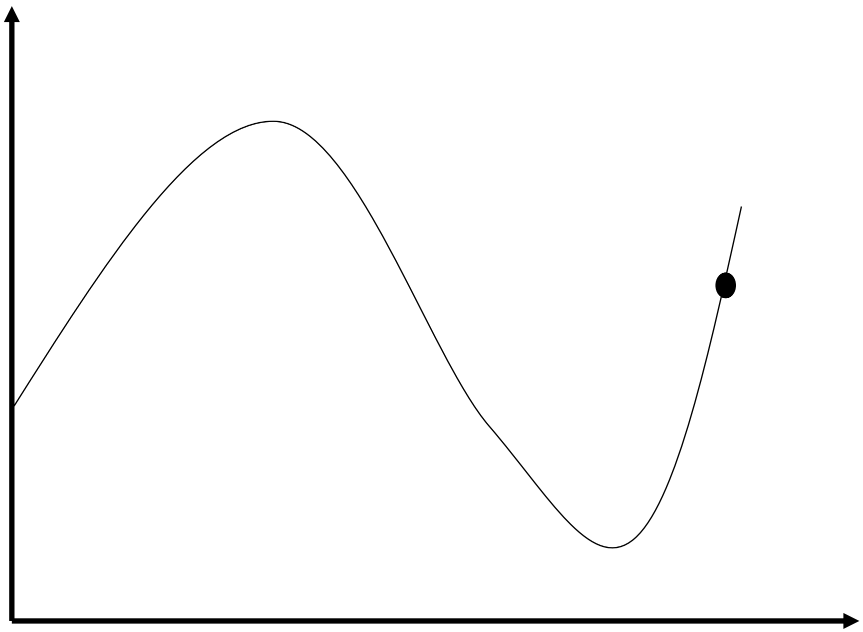


Rule for finding minimum:

- if f′(x) < 0 move right
- if f′(x) > 0 move left
- if f′(x) = 0 stop

Gradient Descent

- Given $f : \mathbb{R} \to \mathbb{R}$



In higher dimensions, we compute $\nabla_x f$

Gradient Descent

- Given $f : \mathbb{R} \to \mathbb{R}$

$$x_{t+1} = x_t + af'(x)$$

- General case $f : \mathbb{R}^n \to \mathbb{R}$

$$x_{t+1} = x_t + a\nabla_x f(x)$$

Value of *a*

- A high learning rate covers more ground at each step, but we risk overshooting the minimum.

- A very low learning rate is more precise but calculating the gradient is time-consuming, so it can take a long time.

Constrained Optimization

- It is the process of optimizing an objective function in the presence of constraints.

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 200 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: extinguish all the fires in minimum time.

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 200 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: extinguish all the fires in minimum time.

- Formulation: Let tA1, tA2, tA3, the times vehicle A devotes to fire 1. ,2, 3.

- Let tB1, tB2, tB3 in a similar way, the times vehicle B devotes to fire 1,2,3.

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 200 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: extinguish all the fires in minimum time.

- Formulation: Let tA1, tA2, tA3, the times vehicle A devotes to fire 1. ,2, 3.

- Let tB1, tB2, tB3 in a similar way, the times vehicle B devotes to fire 1,2,3.

- min T, Constraints?

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 200 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: min T

- Constraints:

1. total units to put out each fire
2. total time should be T

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 2000 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: min T

- Constraints:

1. tA1 + 2tB1 = 1000, tA2 + 2tB2 = 2000, tA3 + 2tB3 = 3000
2. total time should be T

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 2000 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: min T

- Constraints:

1. $tA1 + 2tB1 = 1000$, $tA2 + 2tB2 = 2000$, $tA3 + 2tB3 = 3000$
2. $tA1 + tA2 + tA3 <= T$, $tB1 + tB2 + tB3 <= T$

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 2000 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: min T

- Constraints:

1. $tA1 + 2tB1 = 1000$, $tA2 + 2tB2 = 2000$, $tA3 + 2tB3 = 3000$
2. $tA1 + tA2 + tA3 <= T$, $tB1 + tB2 + tB3 <= T$
3. Other constraints?

Example Problem

- Let Fire 1 need 1000 units of water, Fire 2 need 2000 units of water, Fire 3 3000 units of water. Aircraft A can deliver 1 unit of water per unit time. Aircraft B can deliver 2 units of water per unit time.

- Goal: min T

- Constraints:

1. $tA1 + 2tB1 = 1000$, $tA2 + 2tB2 = 2000$, $tA3 + 2tB3 = 3000$
2. $tA1 + tA2 + tA3 \le T$, $tB1 + tB2 + tB3 \le T$
3. $tAi, tBi, T \ge 0$

Constrained Optimization

- Many problems in engineering can be defined with:
  - a set of constraints defining all candidate (“feasible”) solutions
  $g(x) <= 0$
  - a cost function defining the “quality” of a solution, $f(x)$

Example Problem 2

$$\text{minimize} \sum_{t=0}^{2} u(t)$$

$$\text{subject to}$$

$$x(t+1) = 2x(t) + u(t)$$

$$x(0) = 0$$

$$x(3) = 10$$

$$u(t) \geq 0$$

Example Problem 2

$$\text{minimize} \sum_{t=0}^{2} u(t)$$

subject to

$$x(1) = 2x(0) + u(0)$$
$$x(2) = 2x(1) + u(1)$$
$$x(3) = 2x(2) + u(2)$$
$$u(t) \geq 0$$
$$x(0) = 0$$
$$x(3) = 10$$

Intuition: the earlier you provide the input the better.

Solutiom: u(0) = 2.5

Constrained Optimization

- Many problems in engineering can be defined with:
  - a set of constraints defining all candidate ("feasible") solutions

  g(x) <= 0
  - a cost function defining the "quality" of a solution, f(x)

- If *f* and *g* are affine function of x, these problems are called linear programs.

Linear Program Standard Form

$$\min\ c_1x_1 + c_2x_2 + ... + c_nx_n = z$$

$$s.t.$$

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1$$

$$a_{21}x_2 + a_{22}x_2 + ... + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_m = b_m$$

Linear Program Standard Form

$$\min\ c^T x = z$$

$$s.t.$$

$$Ax = b$$

$$x >= 0$$

- Simplex Algorithm

Intuition Behind Simplex Algorithm

$$\max\ z = x_1 + 2x_2$$

$$s.t.$$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$

Intuition Behind Simplex Algorithm

$$\max\ z = x_1 + 2x_2$$

$$s.t.$$

$$x_1 \le 3$$

$$x_1 + x_2 \le 5$$

$$x_1, x_2 \ge 0$$

Intuition Behind Simplex Algorithm

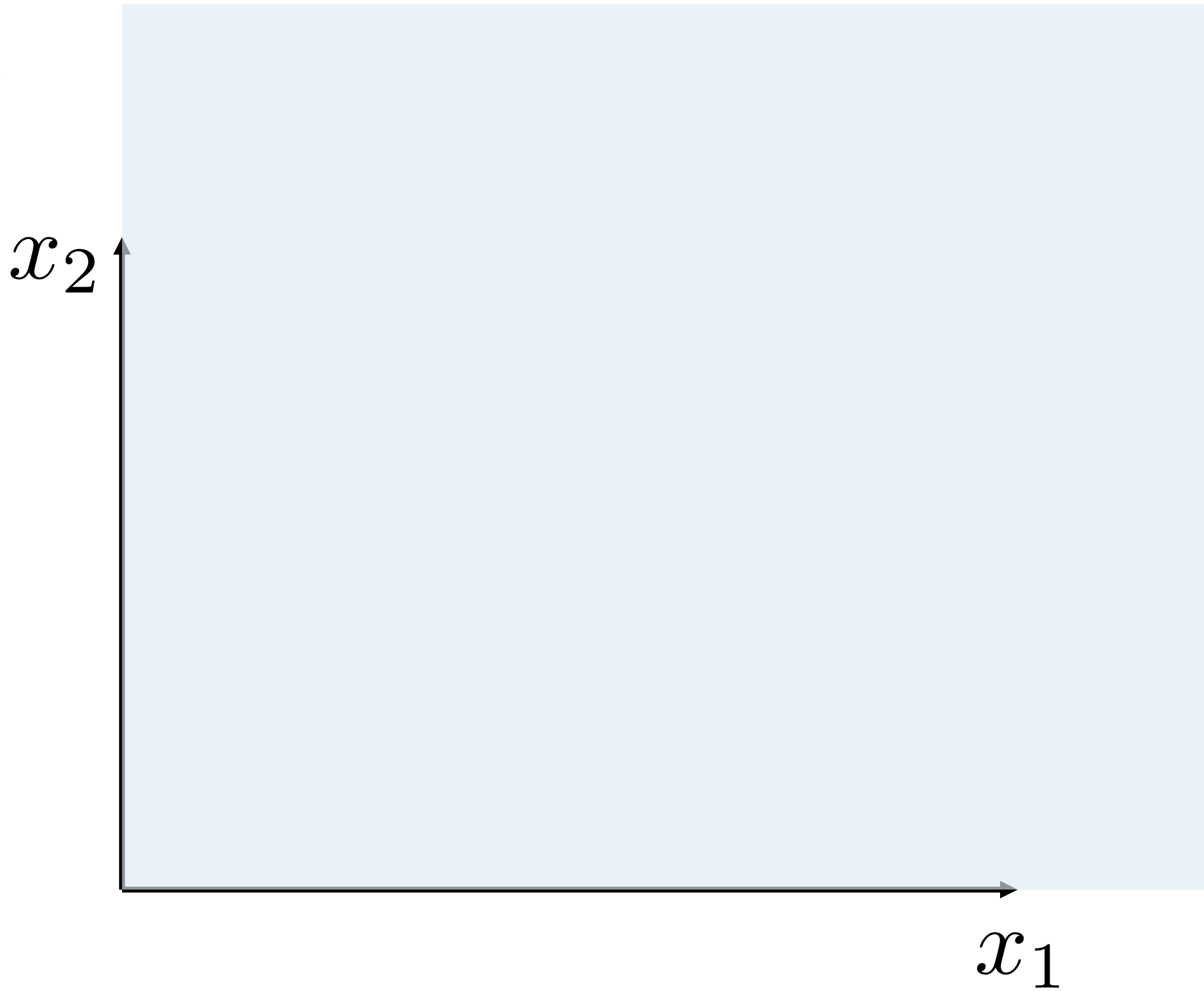
$$\max\ z = x_1 + 2x_2$$

$$s.t.$$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition Behind Simplex Algorithm

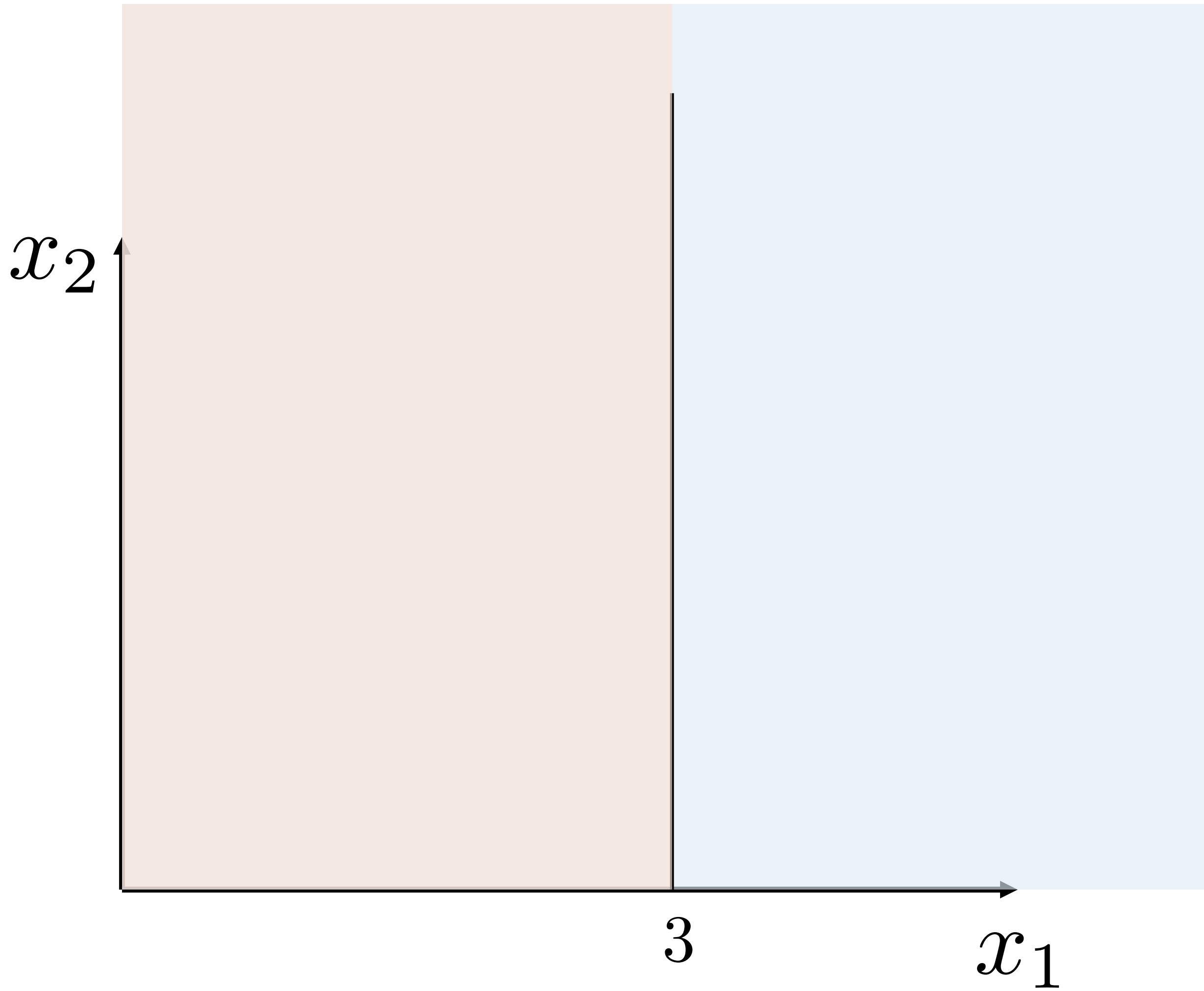
$$\max\ z = x_1 + 2x_2$$

$$s.t.$$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition Behind Simplex Algorithm

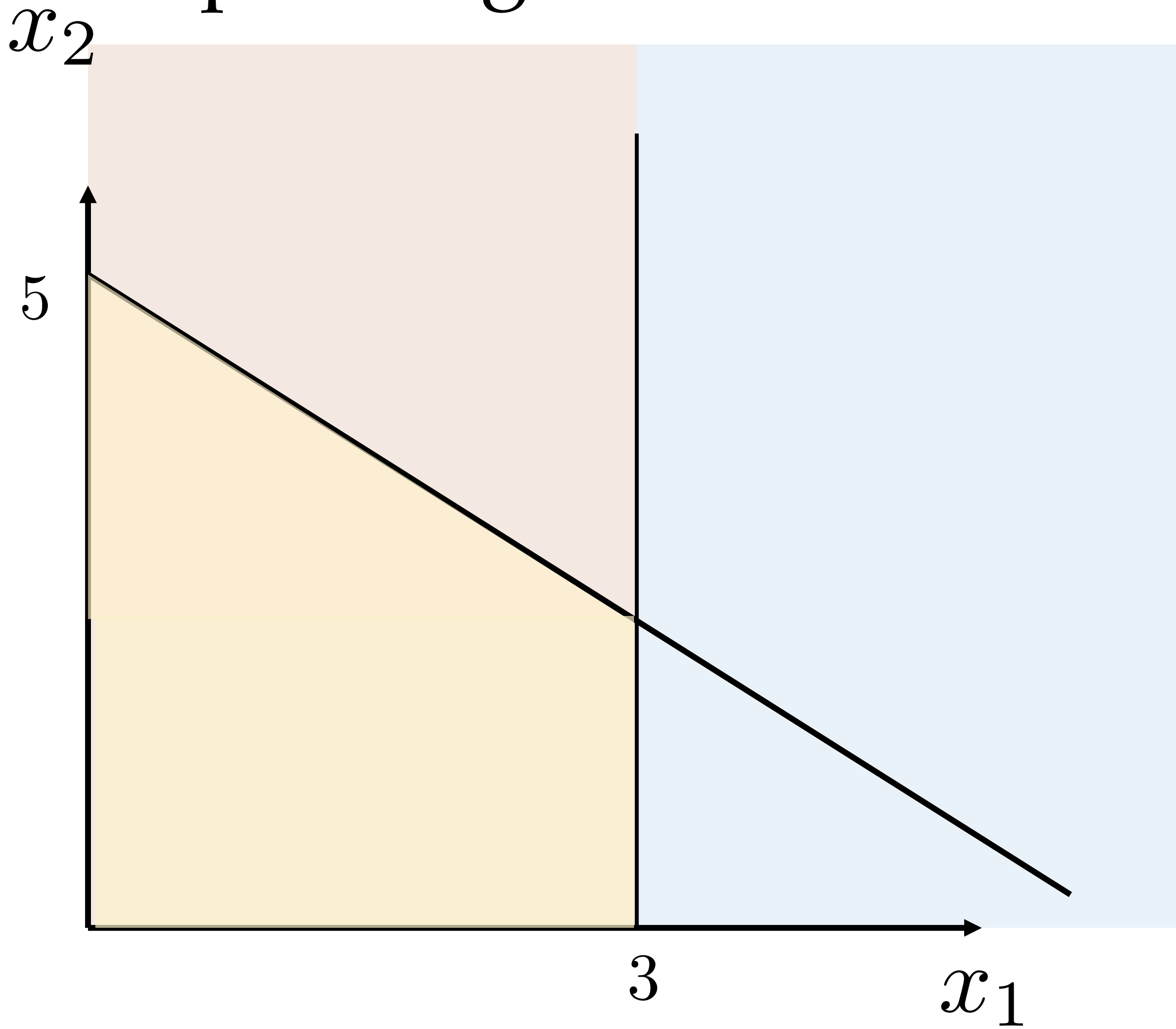
$$\max\ z = x_1 + 2x_2$$

$$s.t.$$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition Behind Simplex Algorithm

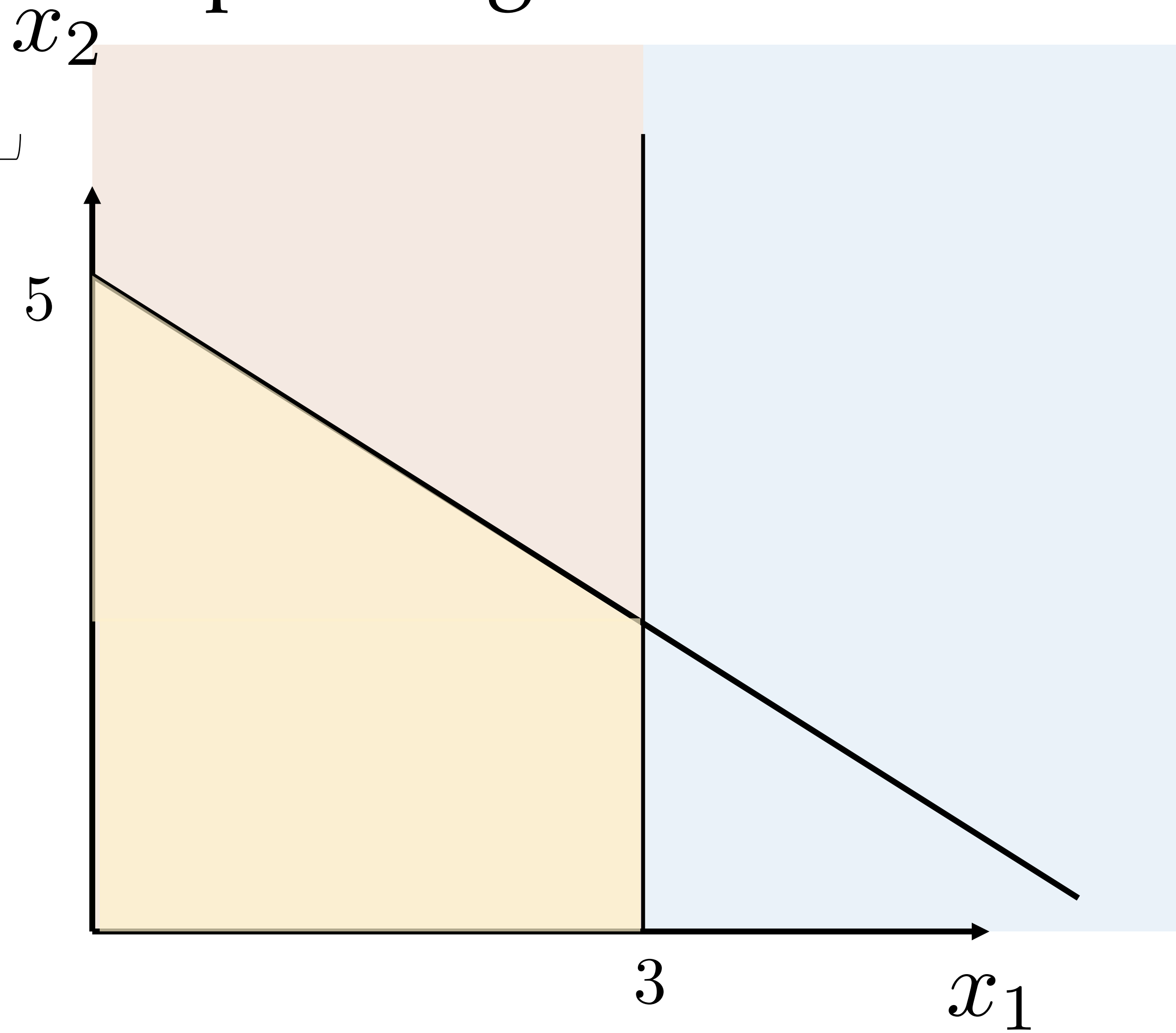
$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1, x_2)}$$

$s.t.$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition Behind Simplex Algorithm

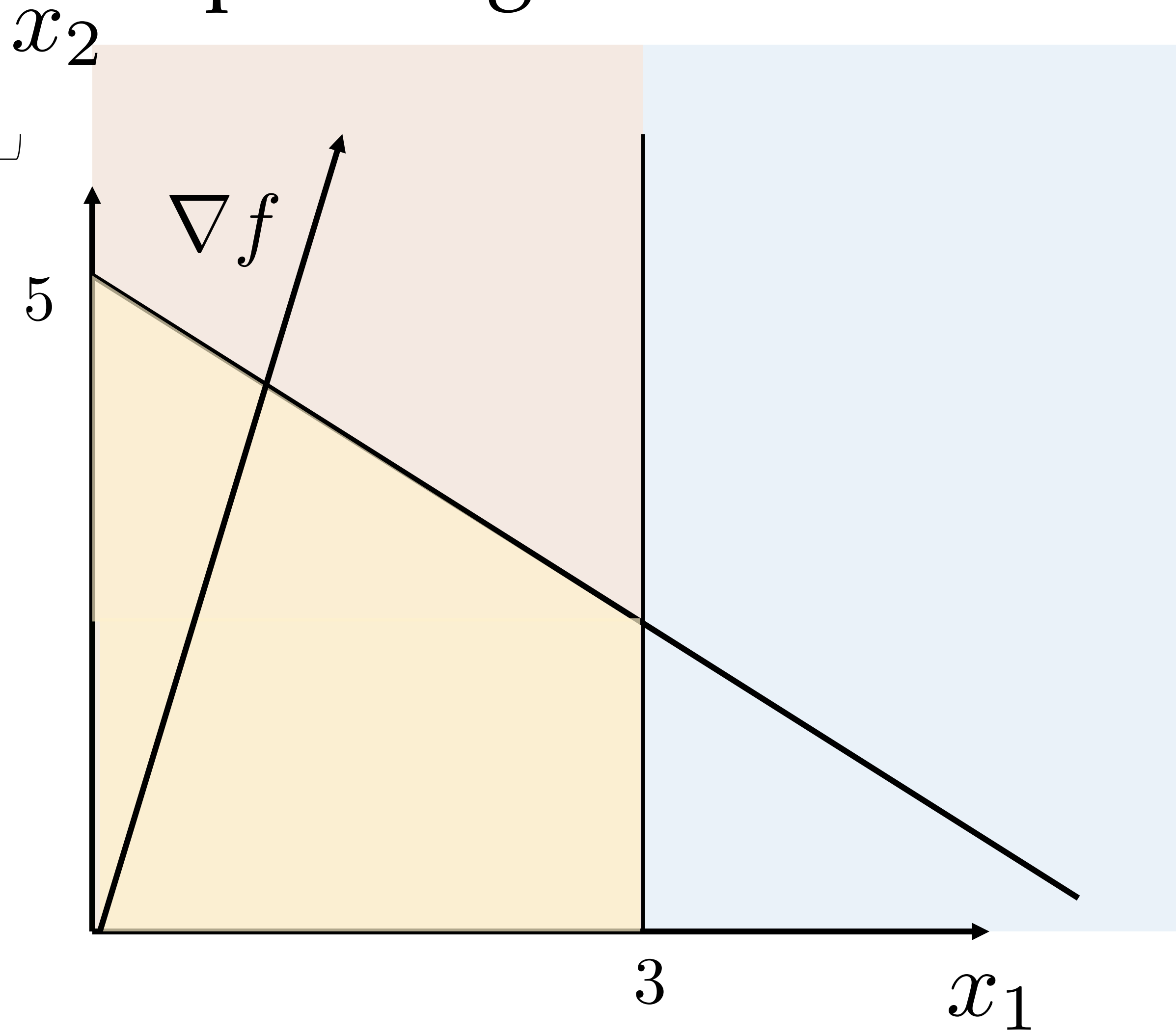
$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1, x_2)}$$

$s.t.$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition Behind Simplex Algorithm

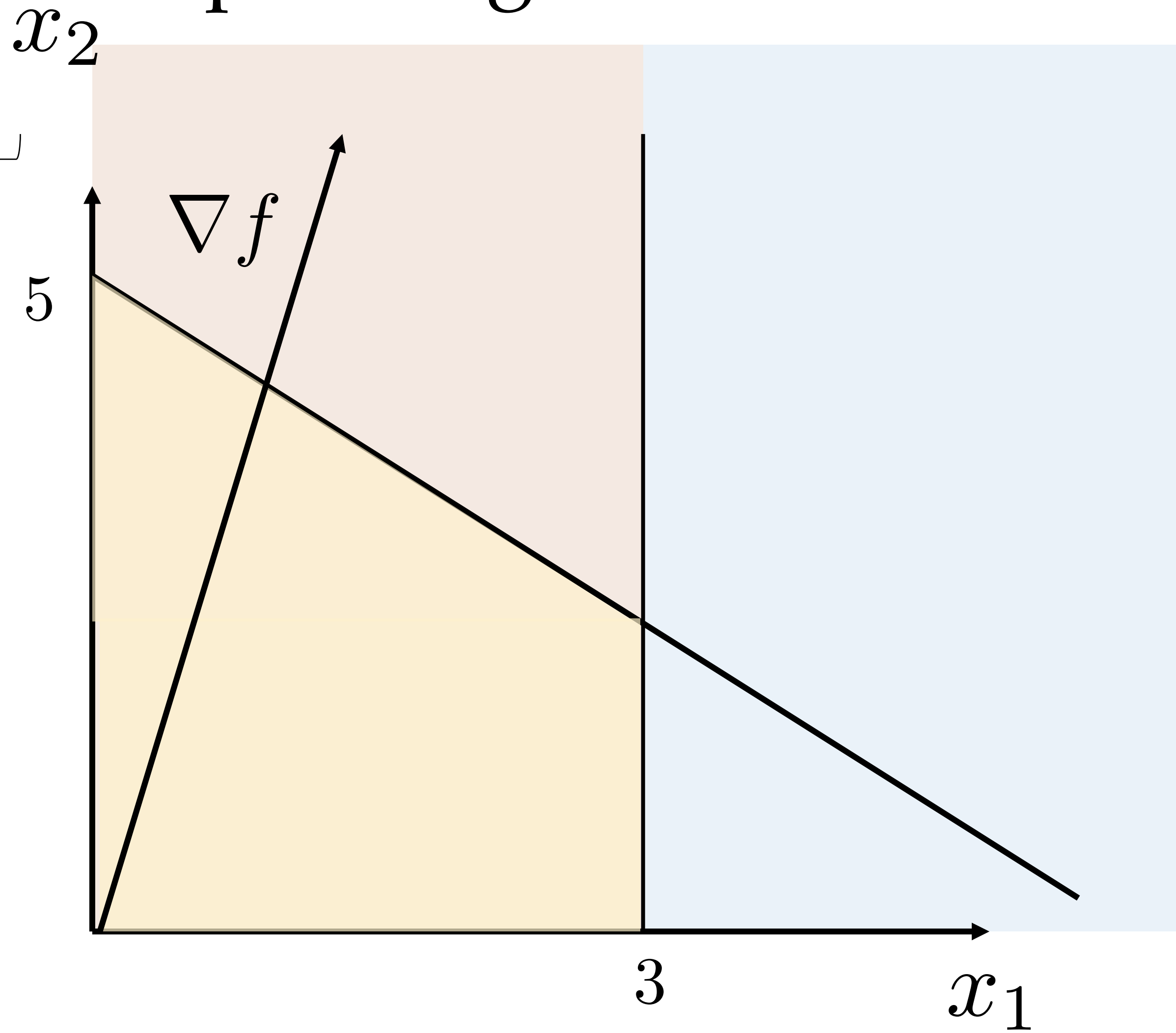
$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1, x_2)}$$

$s.t.$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition: If we are not in a corner, there is always a direction we can go to improve the function.

Intuition Behind Simplex Algorithm

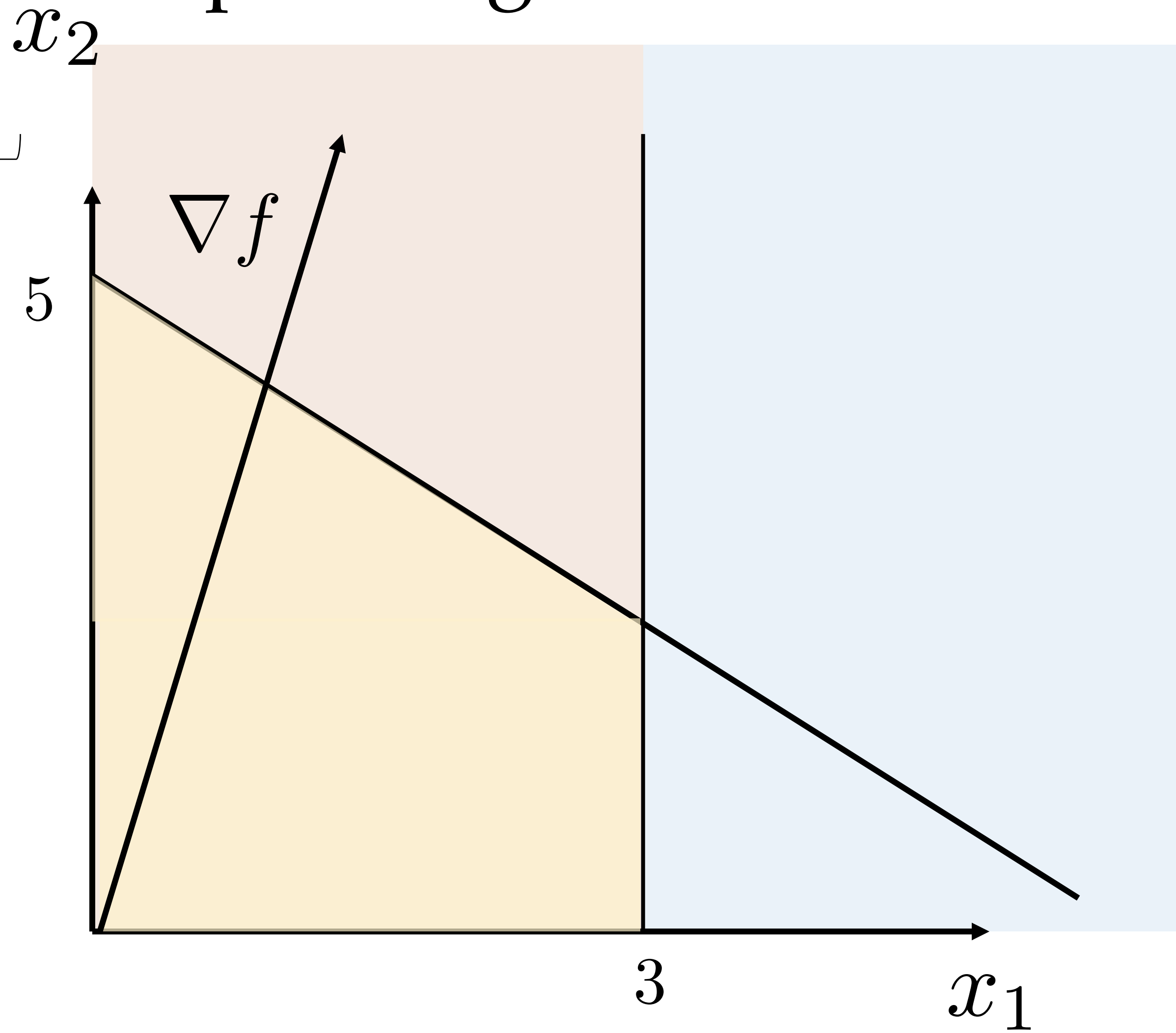
$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1, x_2)}$$

$s.t.$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



Intuition: If we are not in a corner, there is always a direction we can go to improve the function.

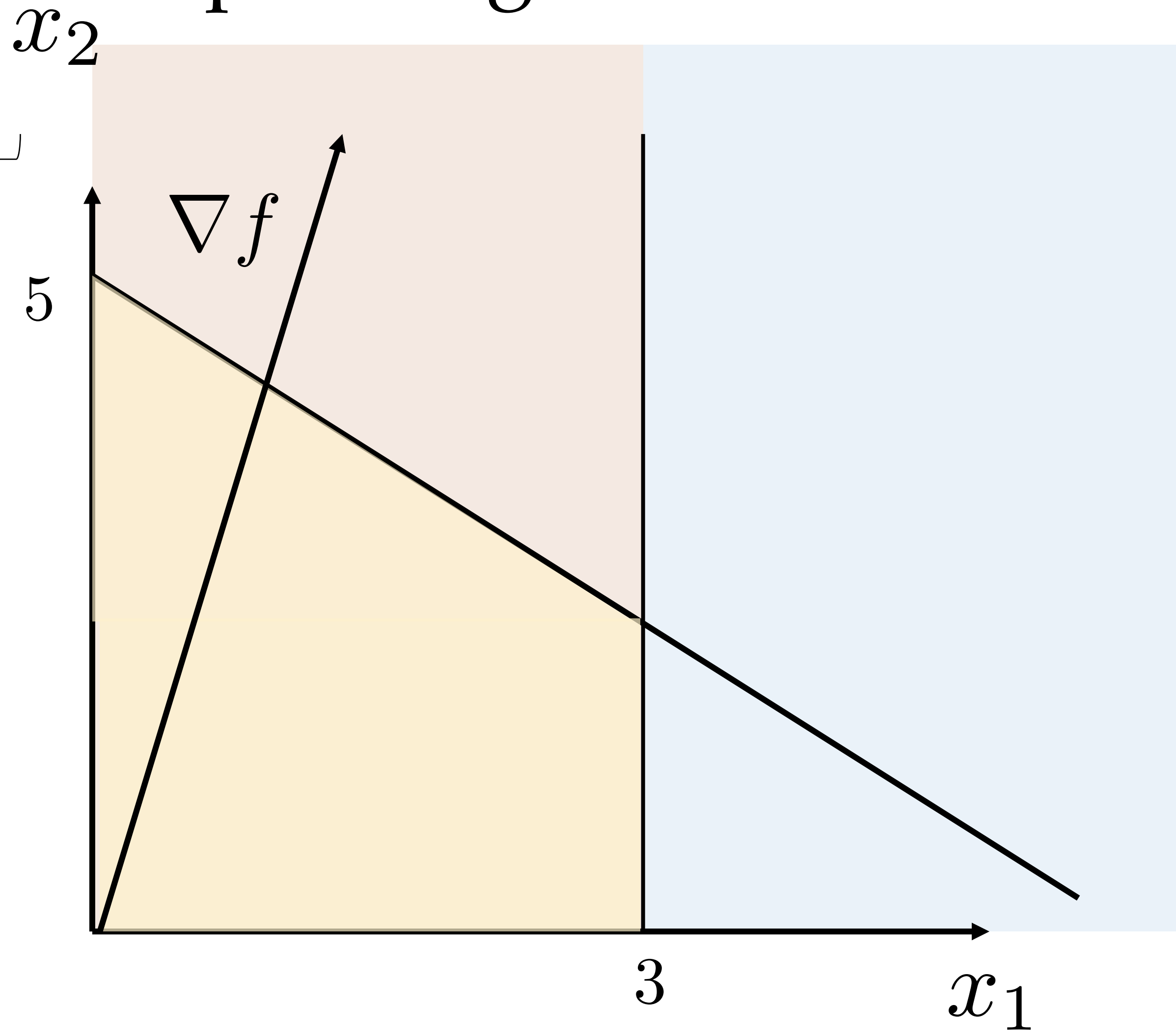
If there is no adjacent corner that improves f, it is the optimum

Intuition Behind Simplex Algorithm

$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1, x_2)}$$

$s.t.$

$$x_1 \leq 3$$
$$x_1 + x_2 \leq 5$$
$$x_1, x_2 \geq 0$$



Intuition: If we are not in a corner, there is always a direction we can go to improve the function.

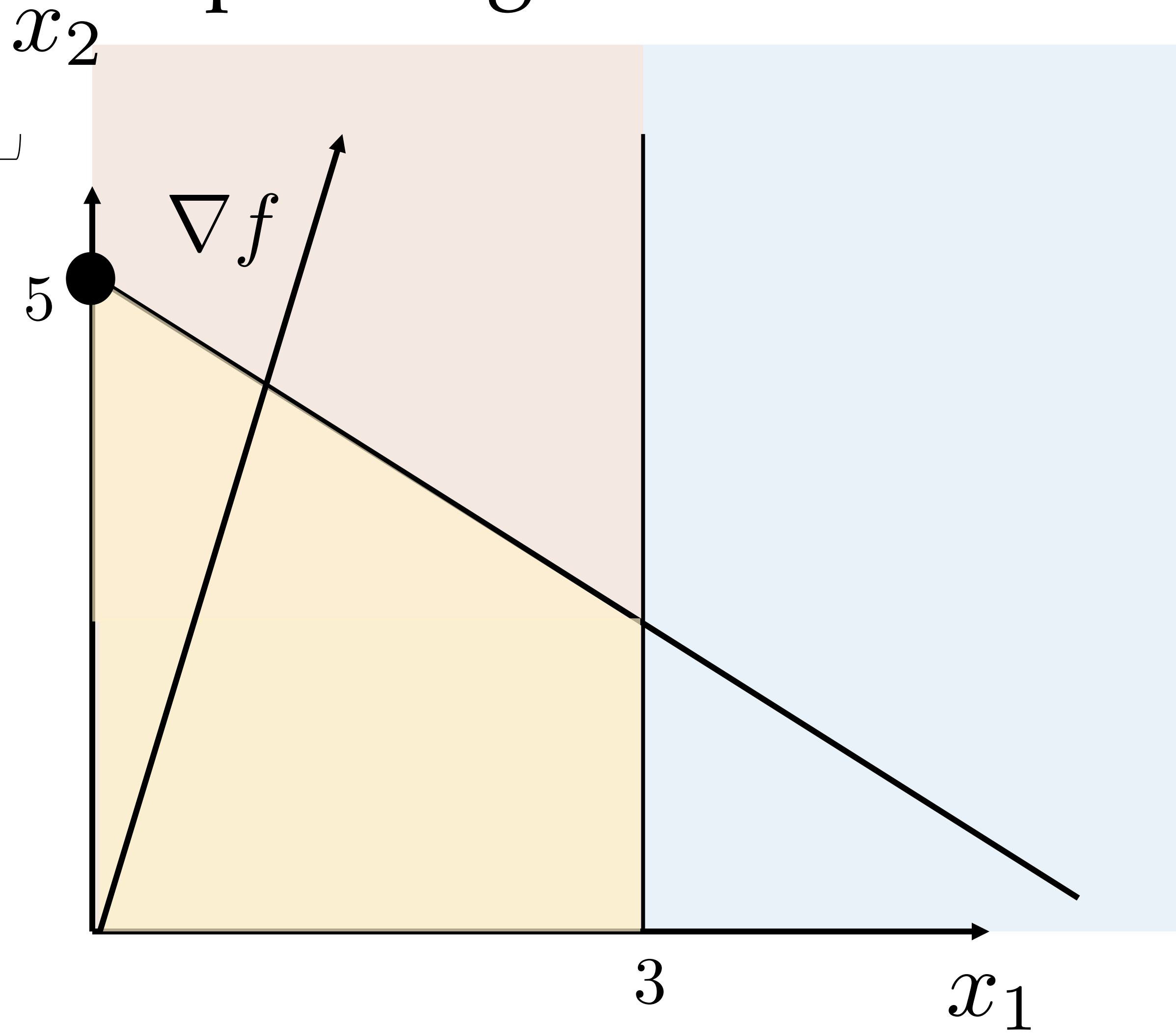
If there is no adjacent corner that improves f, it is the optimum

Intuition Behind Simplex Algorithm

$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1,x_2)}$$

$s.t.$

$$x_1 \leq 3$$
$$x_1 + x_2 \leq 5$$
$$x_1, x_2 \geq 0$$



Intuition: If we are not in a corner, there is always a direction we can go to improve the function.

If there is no adjacent corner that improves f, it is the optimum

Intuition Behind Simplex Algorithm

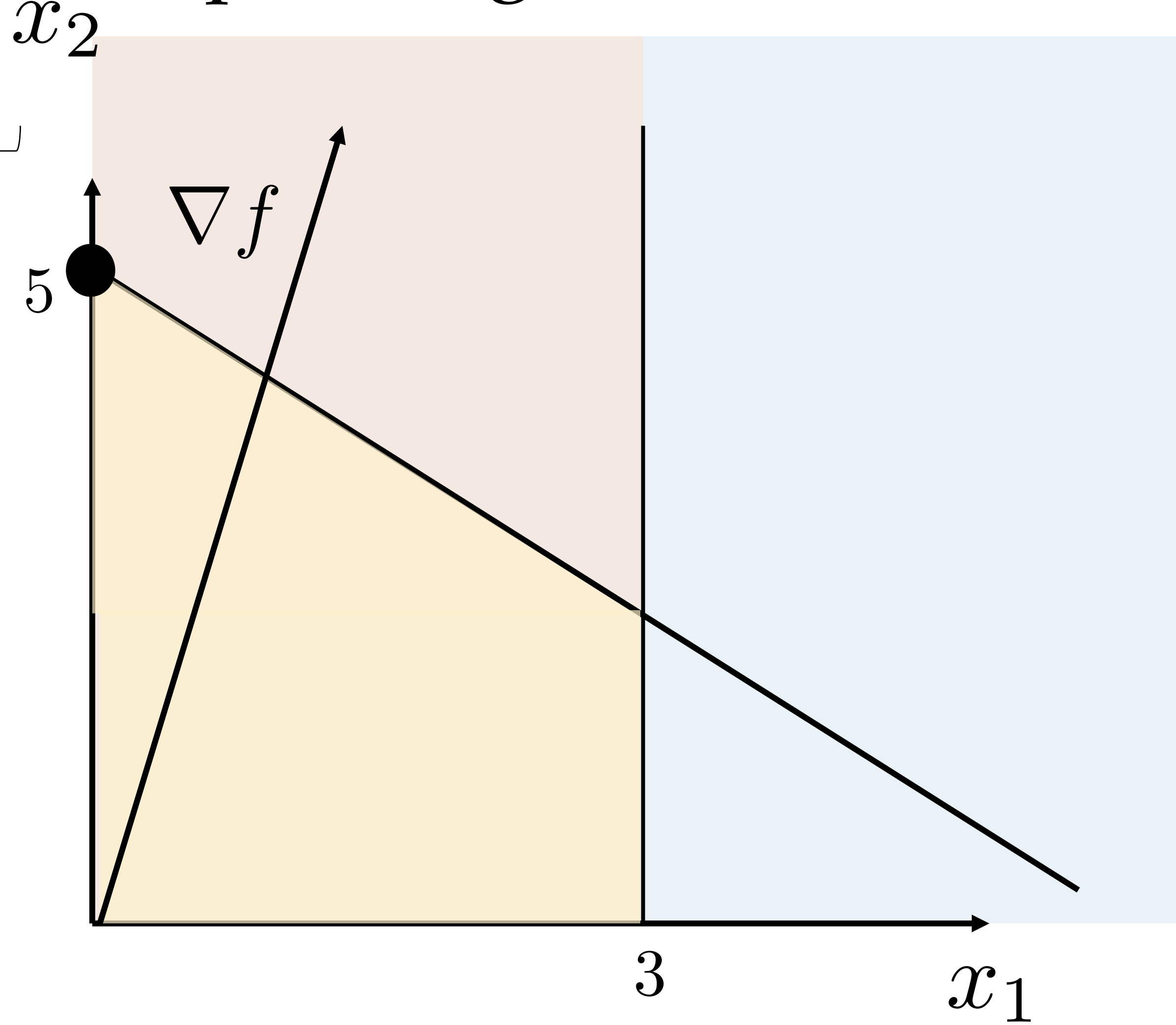
$$\max\ z = \underbrace{x_1 + 2x_2}_{f(x_1, x_2)}$$

$s.t.$

$$x_1 \leq 3$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$



When does this not hold?

Constrained Optimization: General Case

- We have arbitrary non-linear objective and constrain functions, *f*, *g*.

- We use a method called *Lagrange Multipliers*

Example: two variables

- We can represent a constraint as a curve $g(x_1, x_2) = 0$

Example: two variables

- To stay on the curve we need to follow the tangent.

Example: two variables

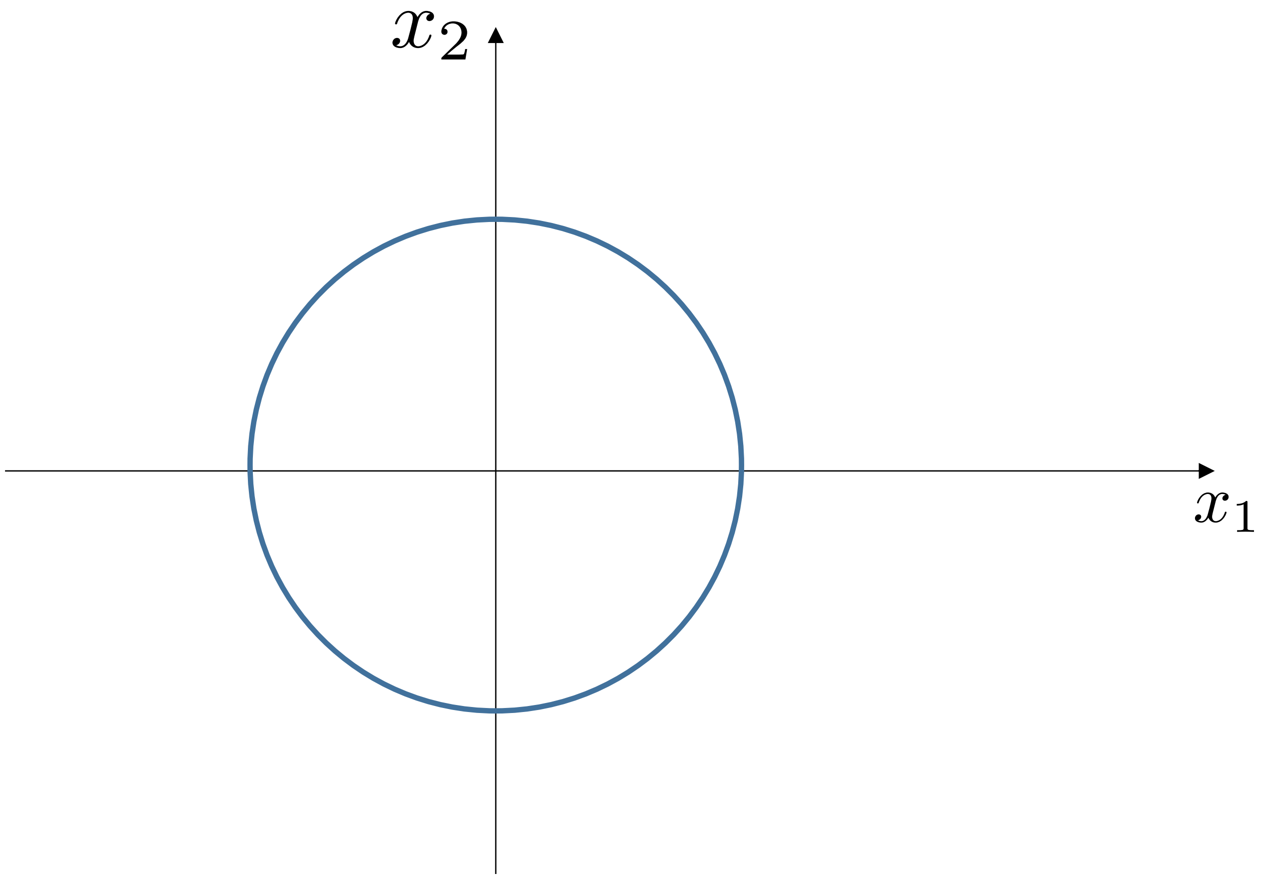
- To stay on the curve we need to follow the tangent.

Example: Circle

$$g(x_1, x_2) = x_1^2 + x_2^2 = 1$$

Example: Circle

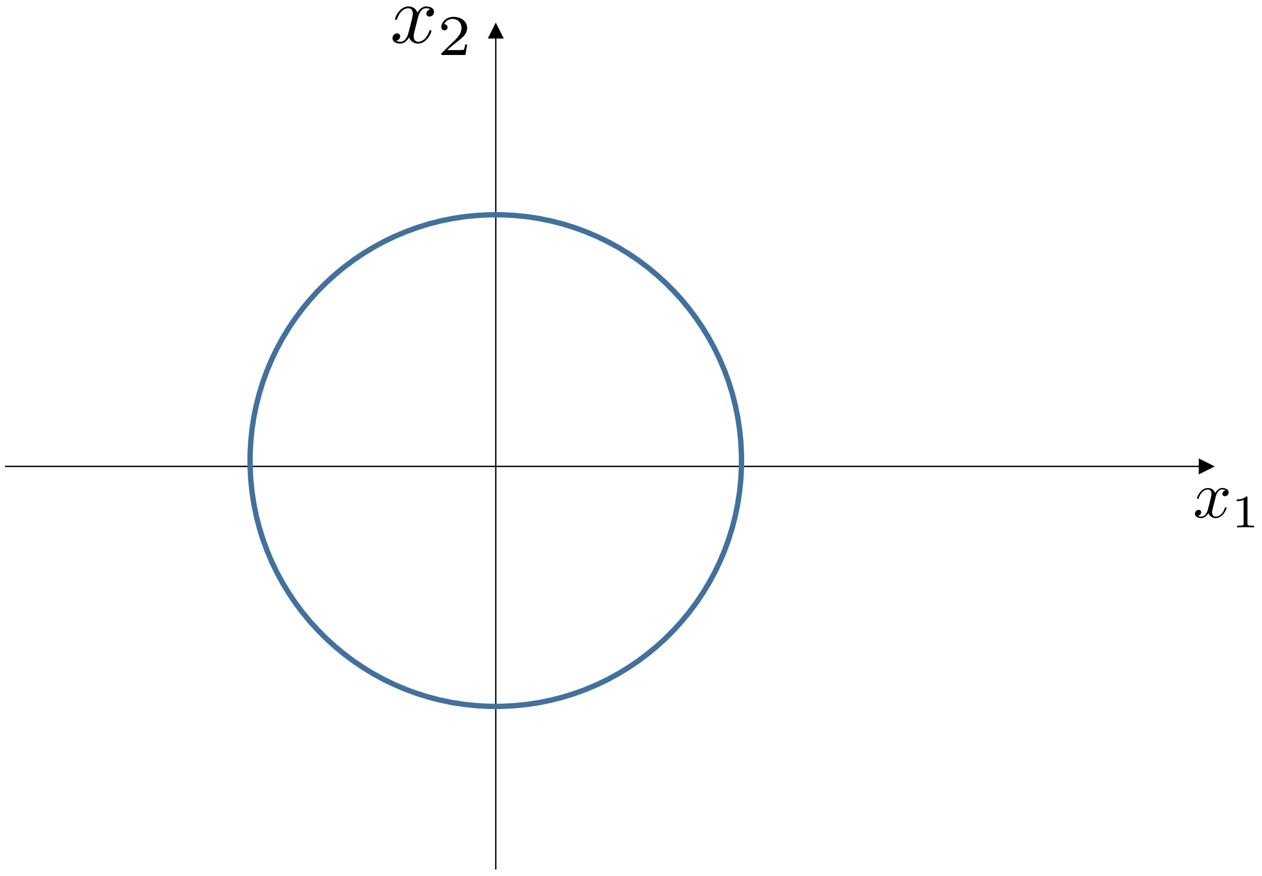
$$g(x_1, x_2) = x_1^2 + x_2^2 = 1$$



Example: Circle

$$g(x_1, x_2) = x_1^2 + x_2^2 = 1$$

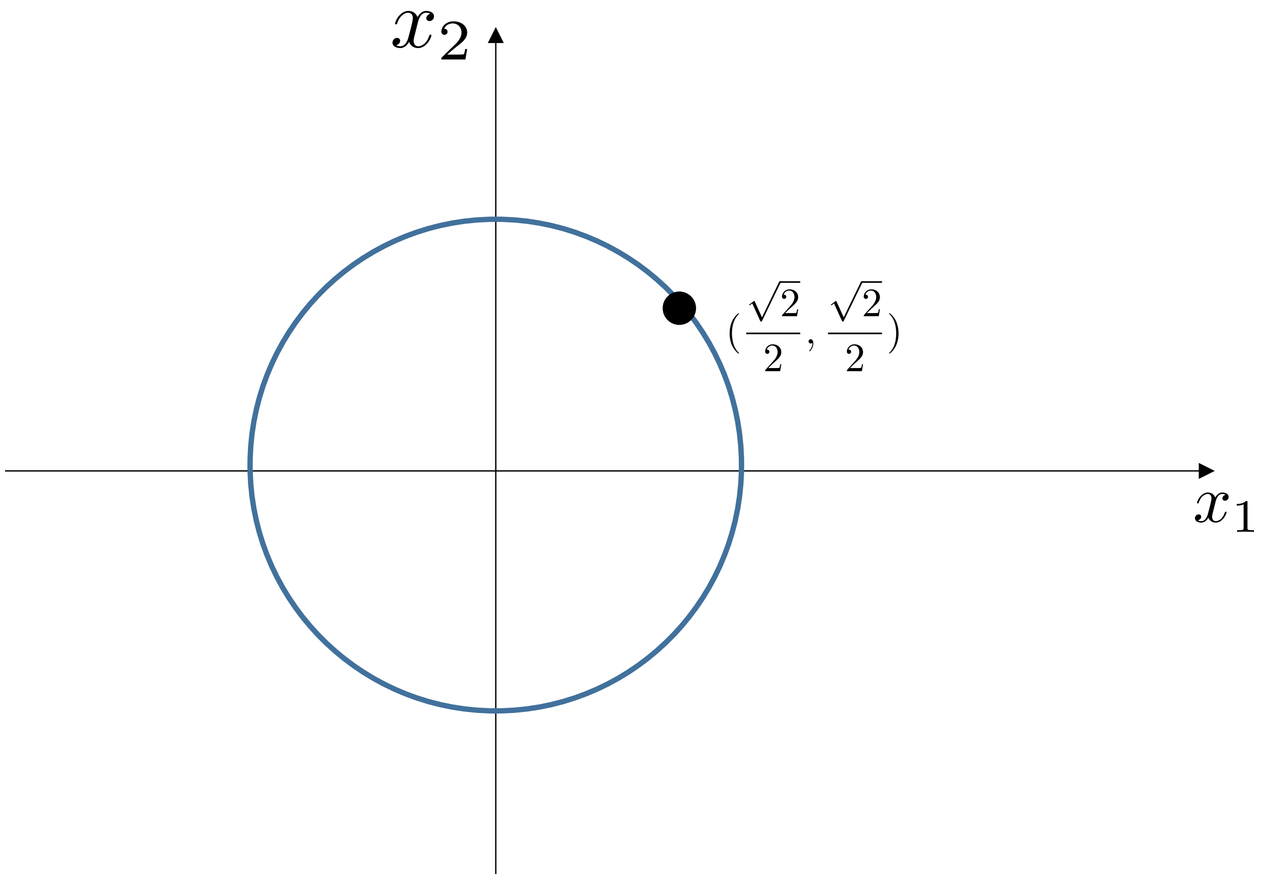
$$\nabla g(x_1, x_2) = (2x_1, 2x_2)$$



Example: Circle

$$g(x_1, x_2) = x_1^2 + x_2^2 = 1$$

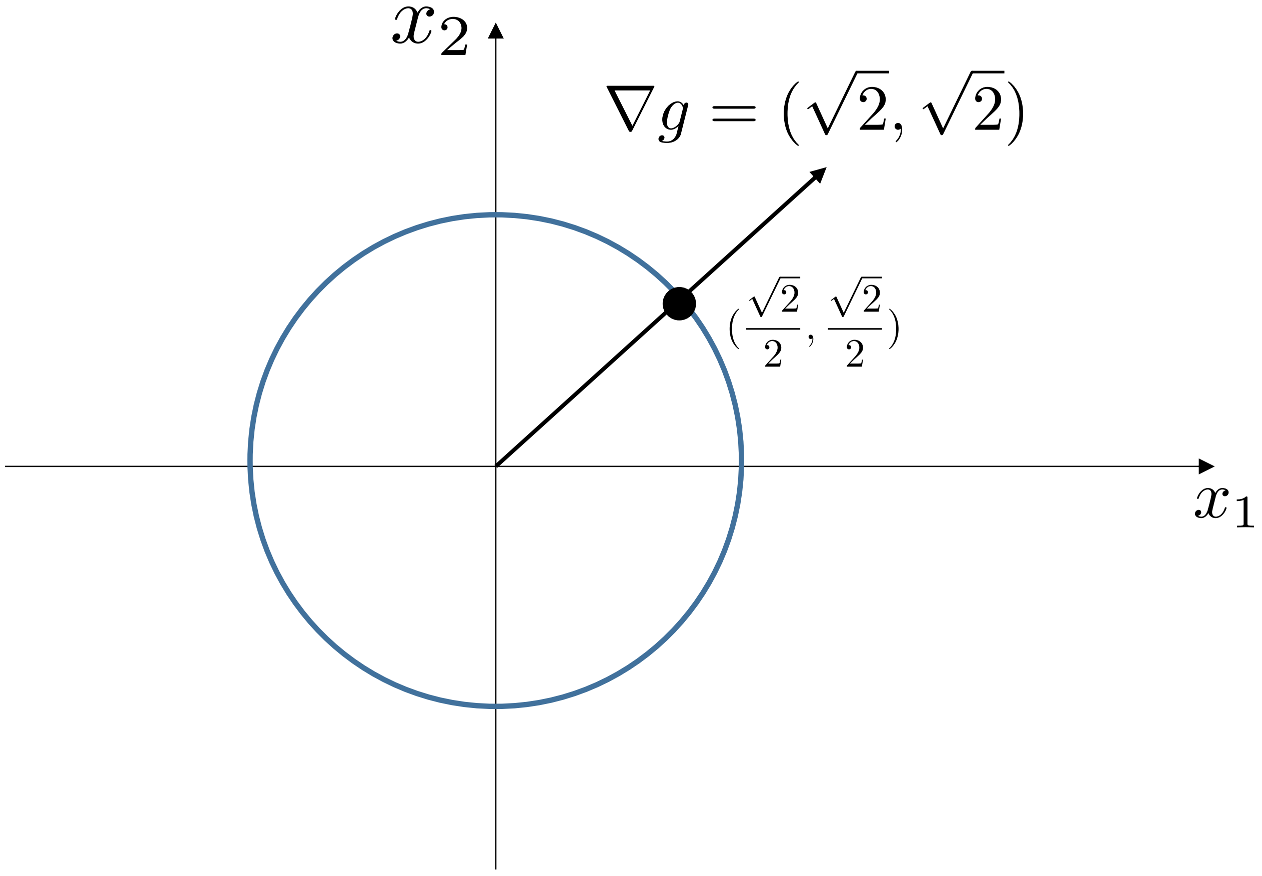
$$\nabla g(x_1, x_2) = (2x_1, 2x_2)$$



Example: Circle

$$g(x_1, x_2) = x_1^2 + x_2^2 = 1$$

$$\nabla g(x_1, x_2) = (2x_1, 2x_2)$$



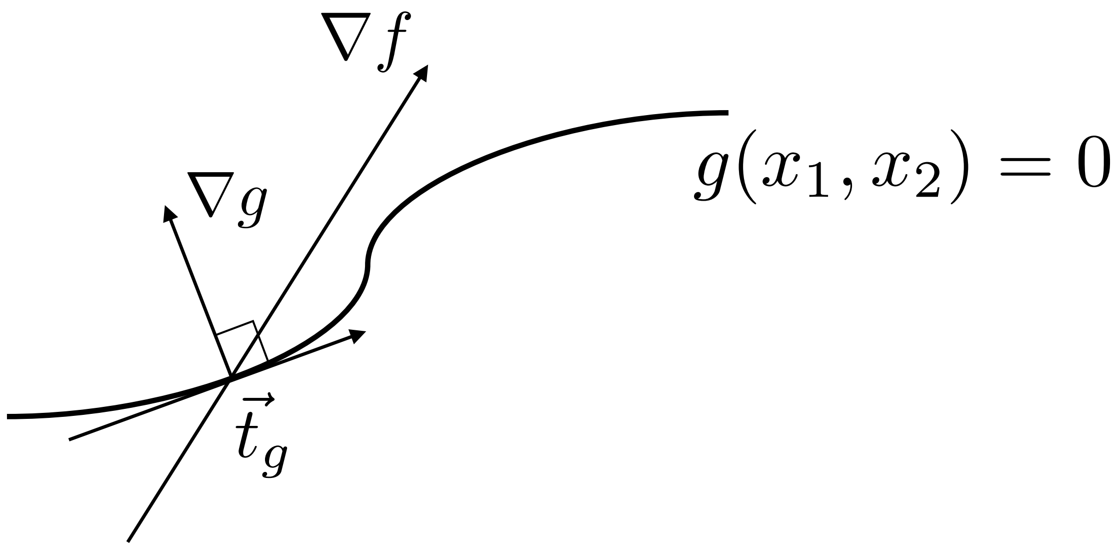
Example: two variables

- To stay on the curve we need to follow the tangent.

$$\nabla g \cdot \vec{t}_g = 0$$

Example: two variables

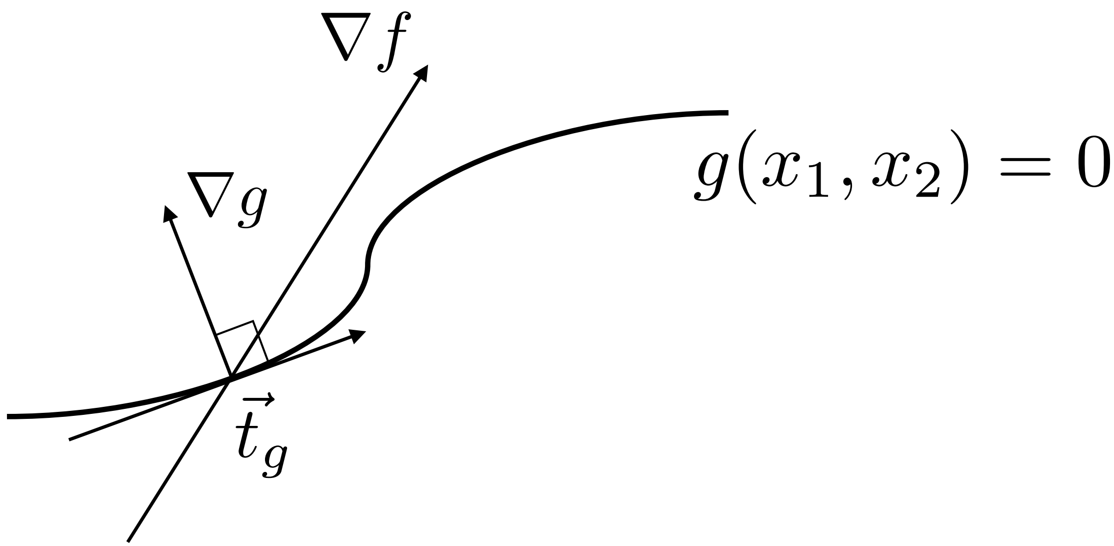
- To stay on the curve we need to follow the tangent.



$$\nabla g \cdot \vec{t}_g = 0$$

Example: two variables

- Generally, motion along the constraint curve will increase or decrease *f*(*x*, *y*)

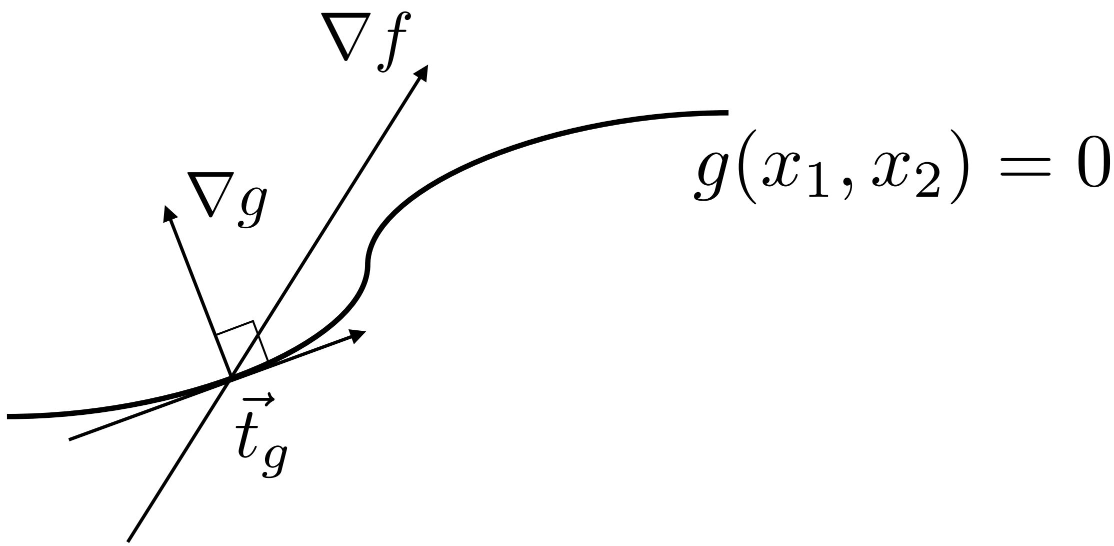


$$\nabla g \cdot \vec{t}_g = 0$$

$$\nabla f \cdot \vec{t}_g \neq 0$$

Example: two variables

- At the optimum:

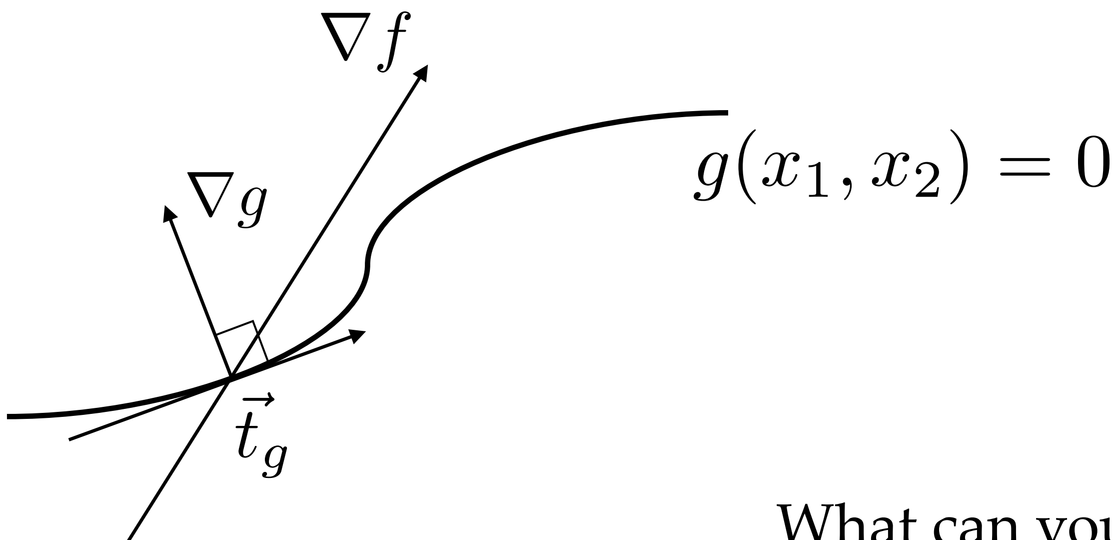


$$\nabla g \cdot \vec{t}_g = 0$$

$$\nabla f \cdot \vec{t}_g = 0$$

Example: two variables

- At the optimum:



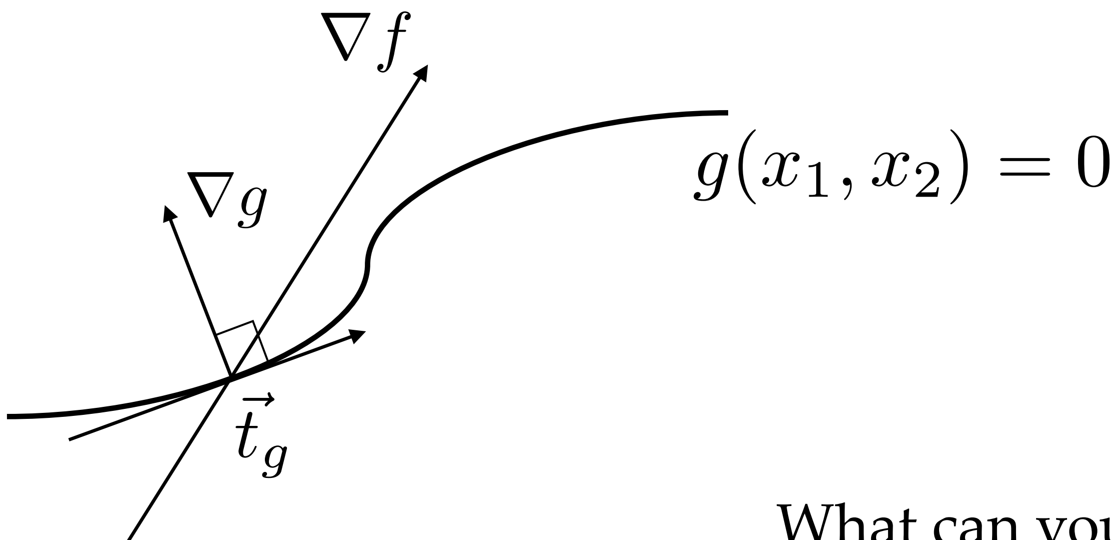
What can you conclude about $\nabla f, \nabla g$?

$$\nabla g \cdot \vec{t}_g = 0$$

$$\nabla f \cdot \vec{t}_g = 0$$

Example: two variables

- At the optimum:



What can you conclude about $\nabla f, \nabla g$?

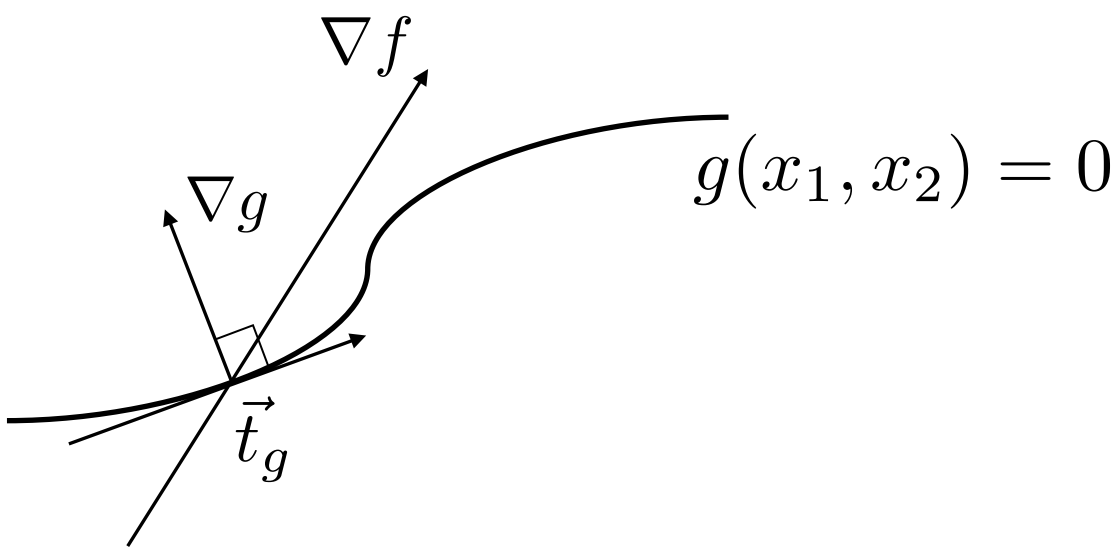
$$\nabla g \cdot \vec{t}_g = 0$$

$$\nabla f \cdot \vec{t}_g = 0$$

$$\nabla f + \lambda \nabla g = 0$$

Example: two variables

- At the optimum:



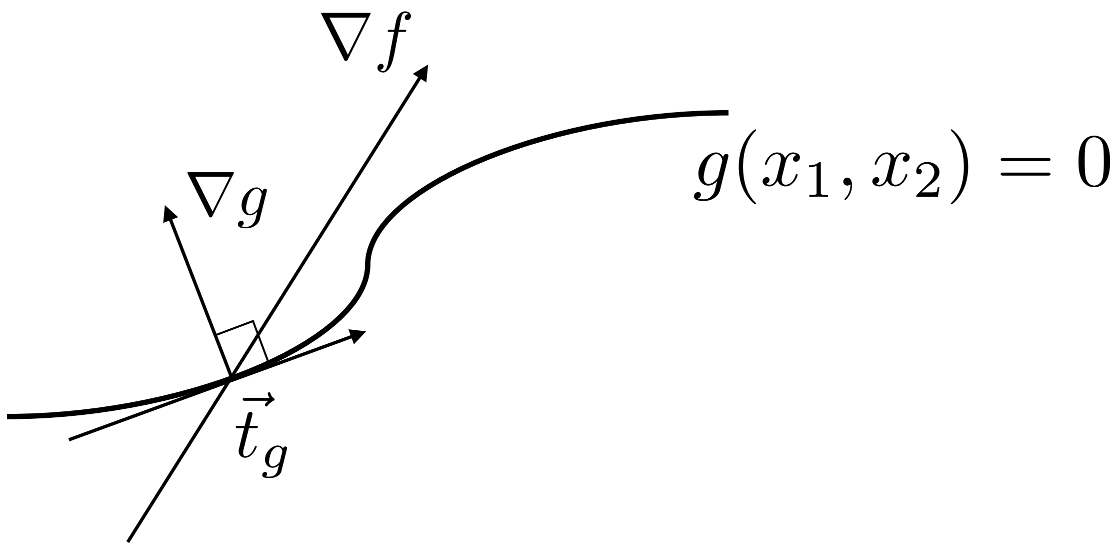
Define the set C_f of points (x_1, x_2) where:

$$g(x_1, x_2) = 0$$

$$\nabla f + \lambda \nabla g = 0$$

Example: two variables

- At the optimum:



Define the set C_f of points (x_1, x_2) where:

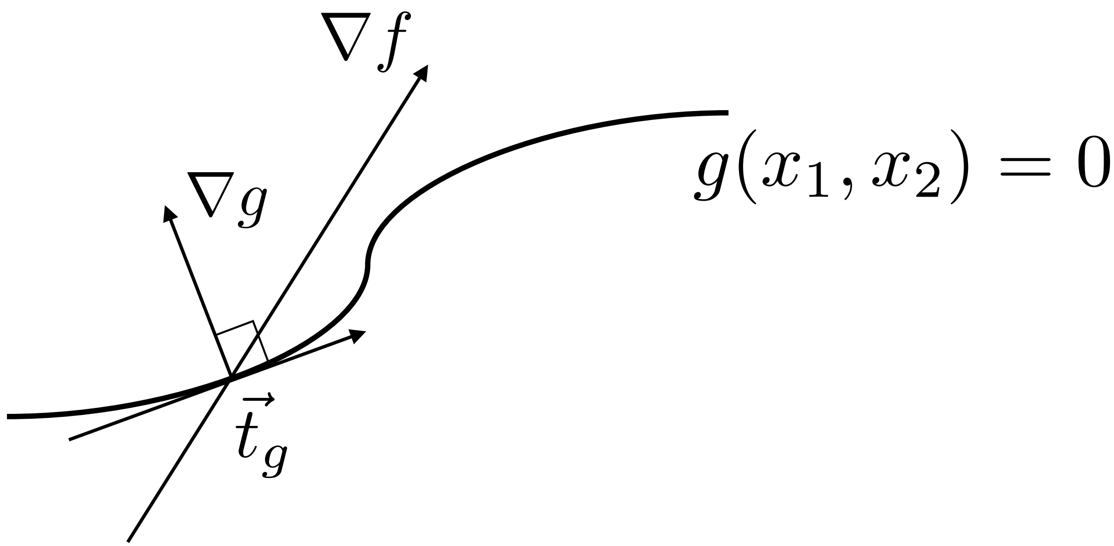
$$g(x_1, x_2) = 0$$

These are the extrema points!

$$\nabla f + \lambda \nabla g = 0$$

Example: two variables

- At the optimum:



Let

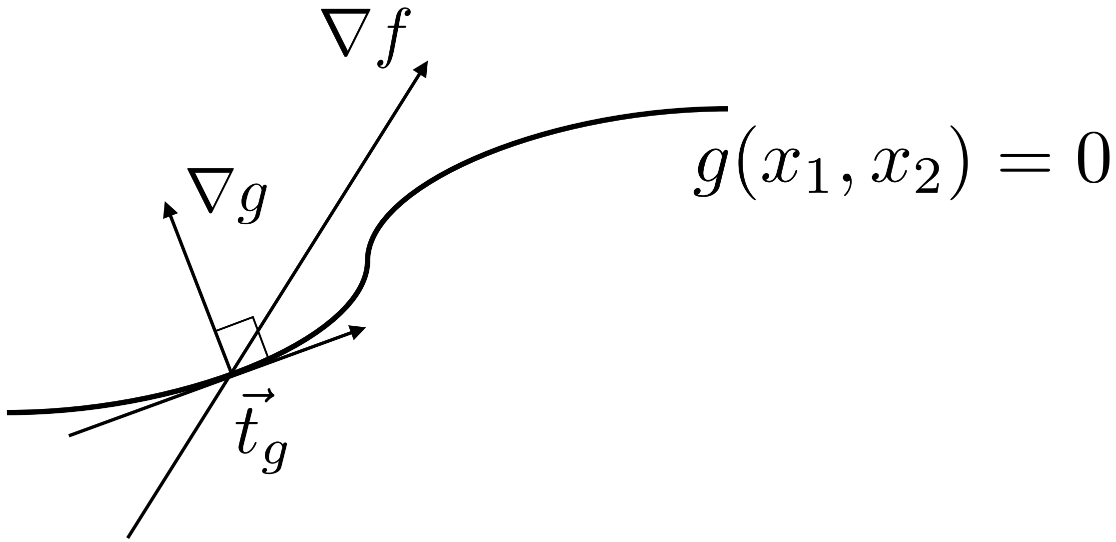
$$F(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

Find x_1, x_2, λ so that:

$$\nabla F = 0$$

Example: two variables

- At the optimum:



Let

$$F(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

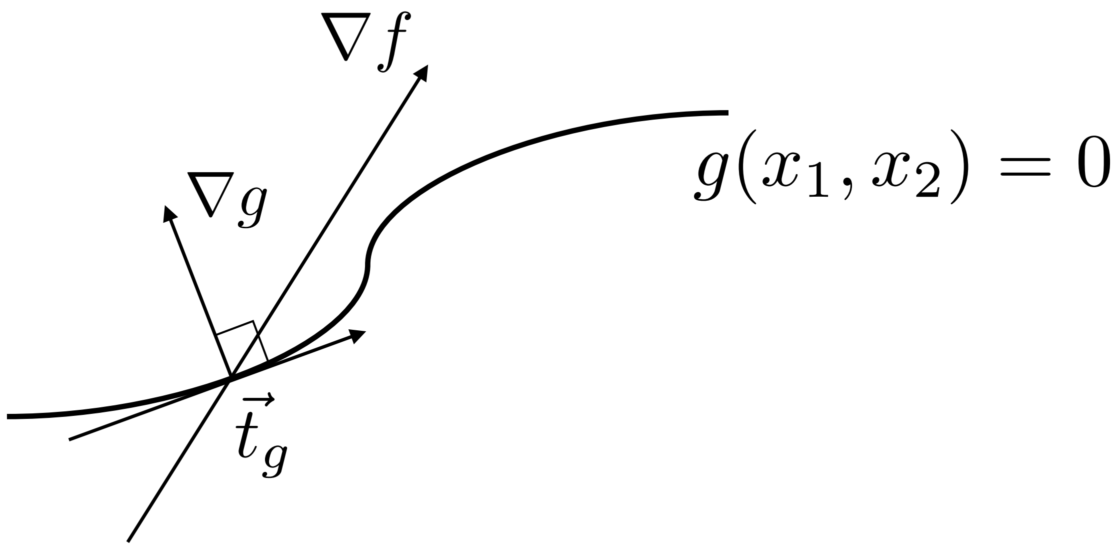
$$\frac{\partial F}{\partial x_1} = 0$$

$$\frac{\partial F}{\partial x_2} = 0$$

$$\frac{\partial F}{\partial \lambda} = 0$$

Example: two variables

- At the optimum:



Let

$$F(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

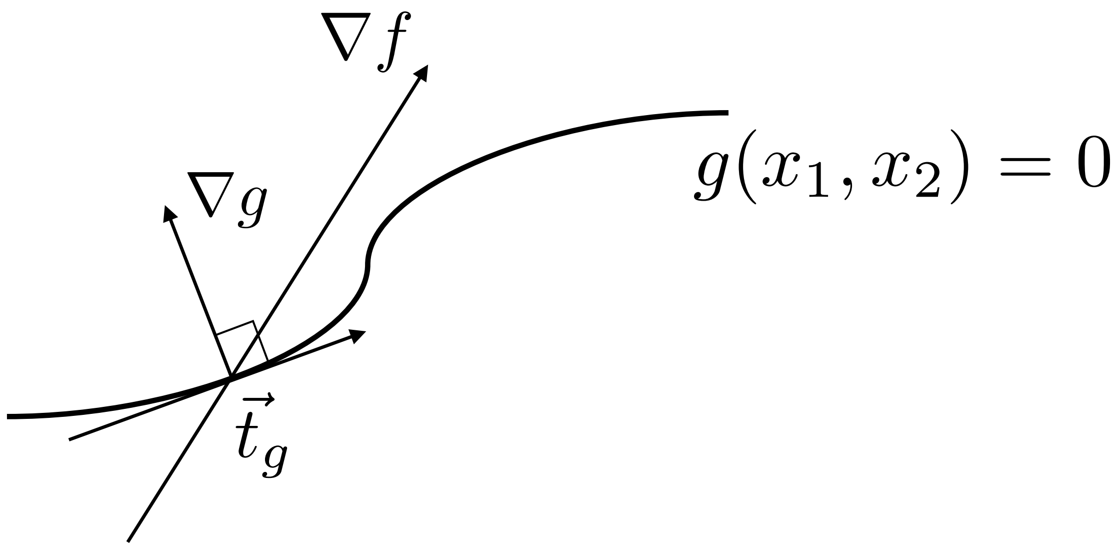
$$\frac{\partial F}{\partial x_1} = \frac{\partial g}{\partial x_1} + \lambda \frac{\partial f}{\partial x_1} = 0$$

$$\frac{\partial F}{\partial x_2} = \frac{\partial g}{\partial x_2} + \lambda \frac{\partial f}{\partial x_2} = 0$$

$$\frac{\partial F}{\partial \lambda} = f(x_1, x_2) = 0$$

Example: two variables

- At the optimum:



Define the set C_f of points (x_1, x_2) where:

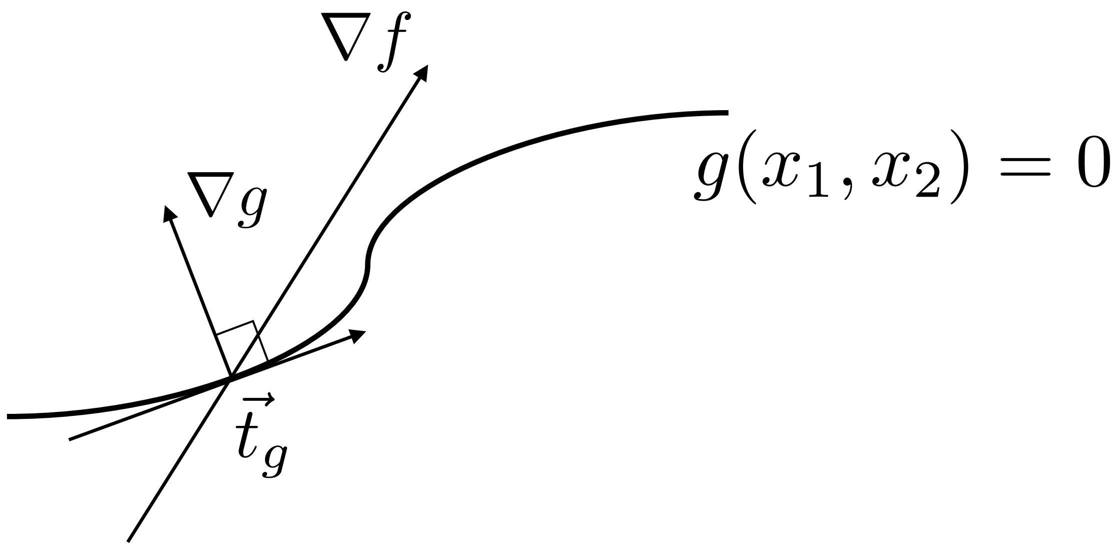
$$g(x_1, x_2) = 0$$

These are the extrema points!

$$\nabla f + \lambda \nabla g = 0$$

Example: two variables

- At the optimum:



Let

$$F(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

$$\frac{\partial F}{\partial x_1} = \frac{\partial g}{\partial x_1} + \lambda \frac{\partial f}{\partial x_1} = 0$$

$$\frac{\partial F}{\partial x_2} = \frac{\partial g}{\partial x_2} + \lambda \frac{\partial f}{\partial x_2} = 0$$

$$\frac{\partial F}{\partial \lambda} = f(x_1, x_2) = 0$$

F: Lagrangian

λ: Lagrange multiplier

Example

- Find the extrema values of the function $f(x_1, x_2) = x_1 x_2$, subject to the constraint:

$$g(x_1, x_2) = \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 = 0$$

$$F(x_1, x_2)?$$

Example

- Find the extrema values of the function $f(x_1, x_2) = x_1 x_2$, subject to the constraint:

$$g(x_1, x_2) = \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 = 0$$

$$F(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

Example

- Find the extrema values of the function $f(x_1, x_2) = x_1 x_2$, subject to the constraint:

$$g(x_1, x_2) = \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 = 0$$

$$F(x_1, x_2) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

Example

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

Example

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

$$\left.\begin{aligned} x_2 &= -\tfrac{\lambda x_1}{4} \\ x_1 &= -\lambda x_2 \\ x_1^2 &= 8 - 4x_2^2 \end{aligned}\right\}$$

Example

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

$$\left.\begin{aligned} x_2 &= -\tfrac{\lambda x_1}{4} \\ x_1 &= -\lambda x_2 \\ x_1^2 &= 8 - 4x_2^2 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} x_2 &= \tfrac{\lambda^2 x_2}{4} \\ x_1 &= -\lambda x_2 \\ x_1^2 &= 8 - 4x_2^2 \end{aligned}\right\}$$

Example

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

$$\left.\begin{aligned} x_2 &= -\tfrac{\lambda x_1}{4} \\ x_1 &= -\lambda x_2 \\ x_1^2 &= 8 - 4x_2^2 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} x_2 &= \tfrac{\lambda^2 x_2}{4} \\ x_1 &= -\lambda x_2 \\ x_1^2 &= 8 - 4x_2^2 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} \lambda &= \pm 2 \\ x_1 &= -\lambda x_2 \\ x_1^2 &= 8 - 4x_2^2 \end{aligned}\right\}$$

Example

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

$$\left.\begin{array}{ll} x_2 & = -\frac{\lambda x_1}{4} \\ x_1 & = -\lambda x_2 \\ x_1^2 & = 8 - 4x_2^2 \end{array}\right\} \Rightarrow \left.\begin{array}{ll} x_2 & = \frac{\lambda^2 x_2}{4} \\ x_1 & = -\lambda x_2 \\ x_1^2 & = 8 - 4x_2^2 \end{array}\right\} \Rightarrow \left.\begin{array}{ll} \lambda & = \pm 2 \\ x_1 & = -\lambda x_2 \\ x_1^2 & = 8 - 4x_2^2 \end{array}\right\} \Rightarrow \left.\begin{array}{ll} \lambda & = \pm 2 \\ x_1 & = \pm 2 x_2 \\ x_2 & = \pm 1 \end{array}\right\}$$

Solution

$$\nabla F(x_1, x_2, \lambda) = \begin{bmatrix} x_2 + \frac{\lambda x_1}{4} \\ x_1 + \lambda x_2 \\ \frac{x_1^2}{8} + \frac{x_2^2}{2} - 1 \end{bmatrix} = 0$$

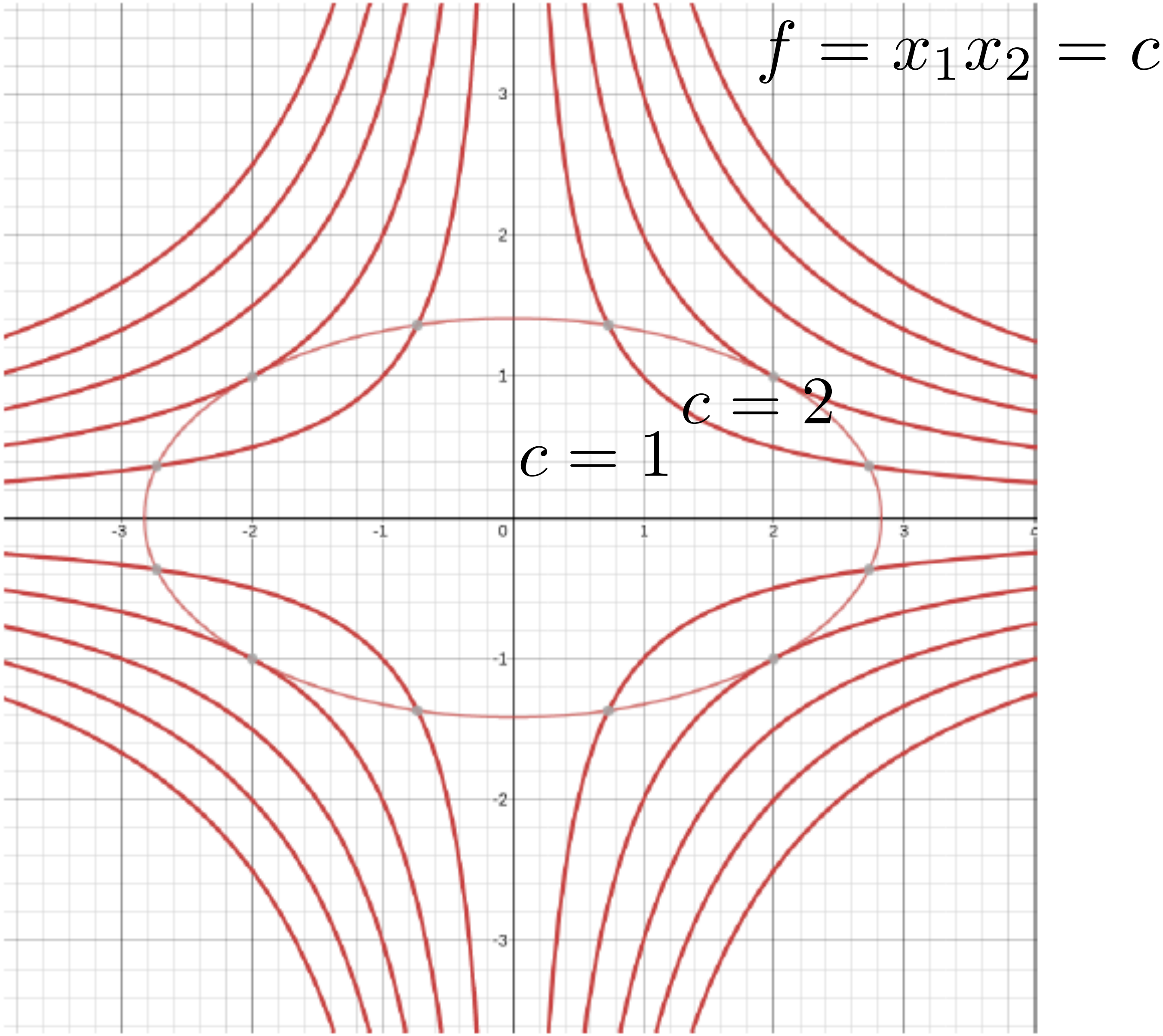
$$\left.\begin{array}{ll} x_2 & = -\frac{\lambda x_1}{4} \\ x_1 & = -\lambda x_2 \\ x_1^2 & = 8 - 4x_2^2 \end{array}\right\} \Rightarrow \left.\begin{array}{ll} x_2 & = \frac{\lambda^2 x_2}{4} \\ x_1 & = -\lambda x_2 \\ x_1^2 & = 8 - 4x_2^2 \end{array}\right\} \Rightarrow \left.\begin{array}{ll} \lambda & = \pm 2 \\ x_1 & = -\lambda x_2 \\ x_1^2 & = 8 - 4x_2^2 \end{array}\right\} \Rightarrow \left.\begin{array}{ll} \lambda & = \pm 2 \\ x_1 & = \pm 2x_2 \\ x_2 & = \pm 1 \end{array}\right\}$$

$$\lambda = \pm 2$$

$$x_1 = \pm 2$$

$$x_2 = \pm 1$$

Solution



Application: Inverse Kinematics

- Given a desired position of the robot's end-effector, find the values of the joint angles that are within their limits and avoid collisions.