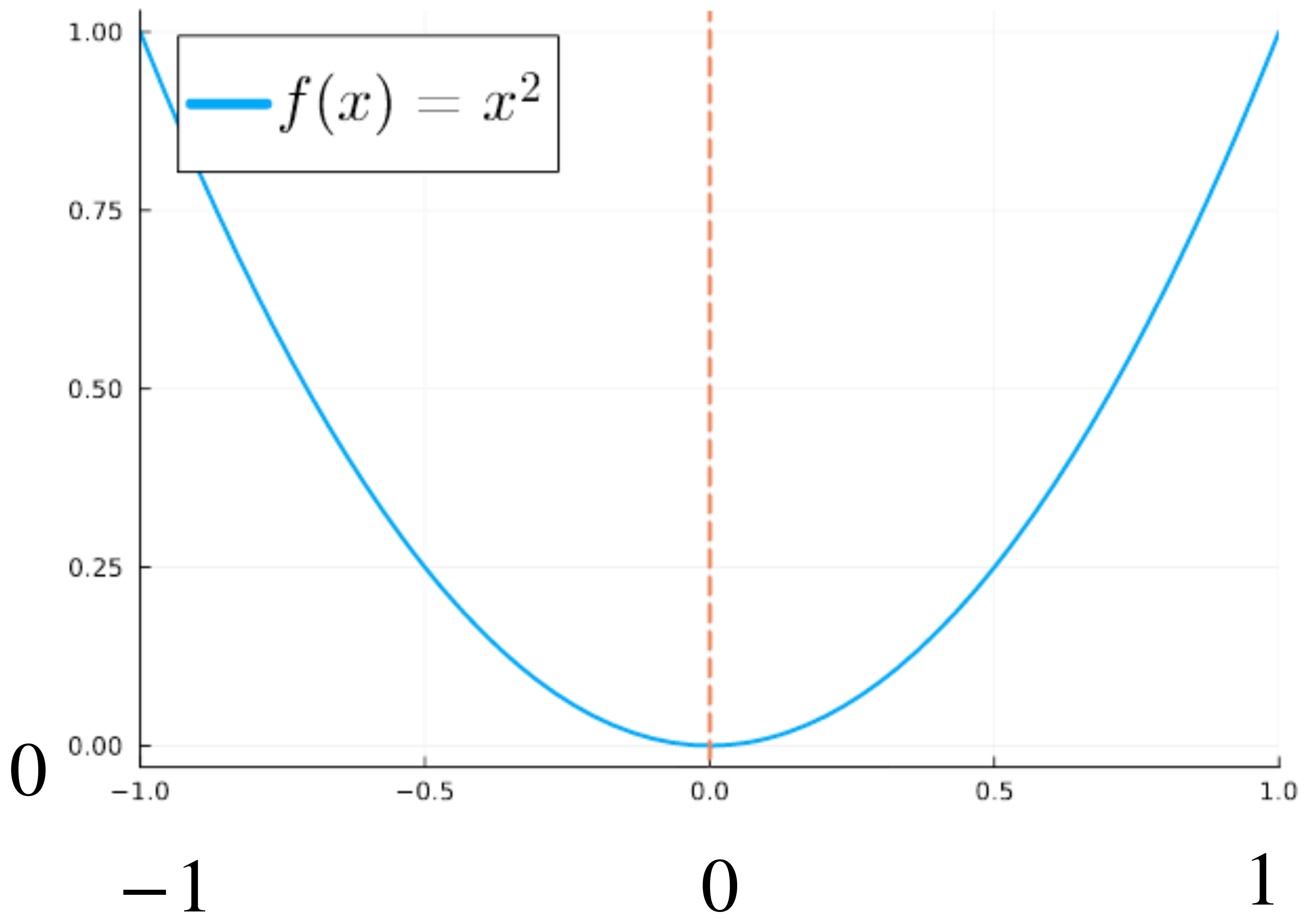
Vector calculus and friends

Algorithmic approaches to maths 2025

Dhruva Venkita Raman

Unconstrained optimisation



Objective

$$\arg\min_x f(x)$$

Derivative at solution?

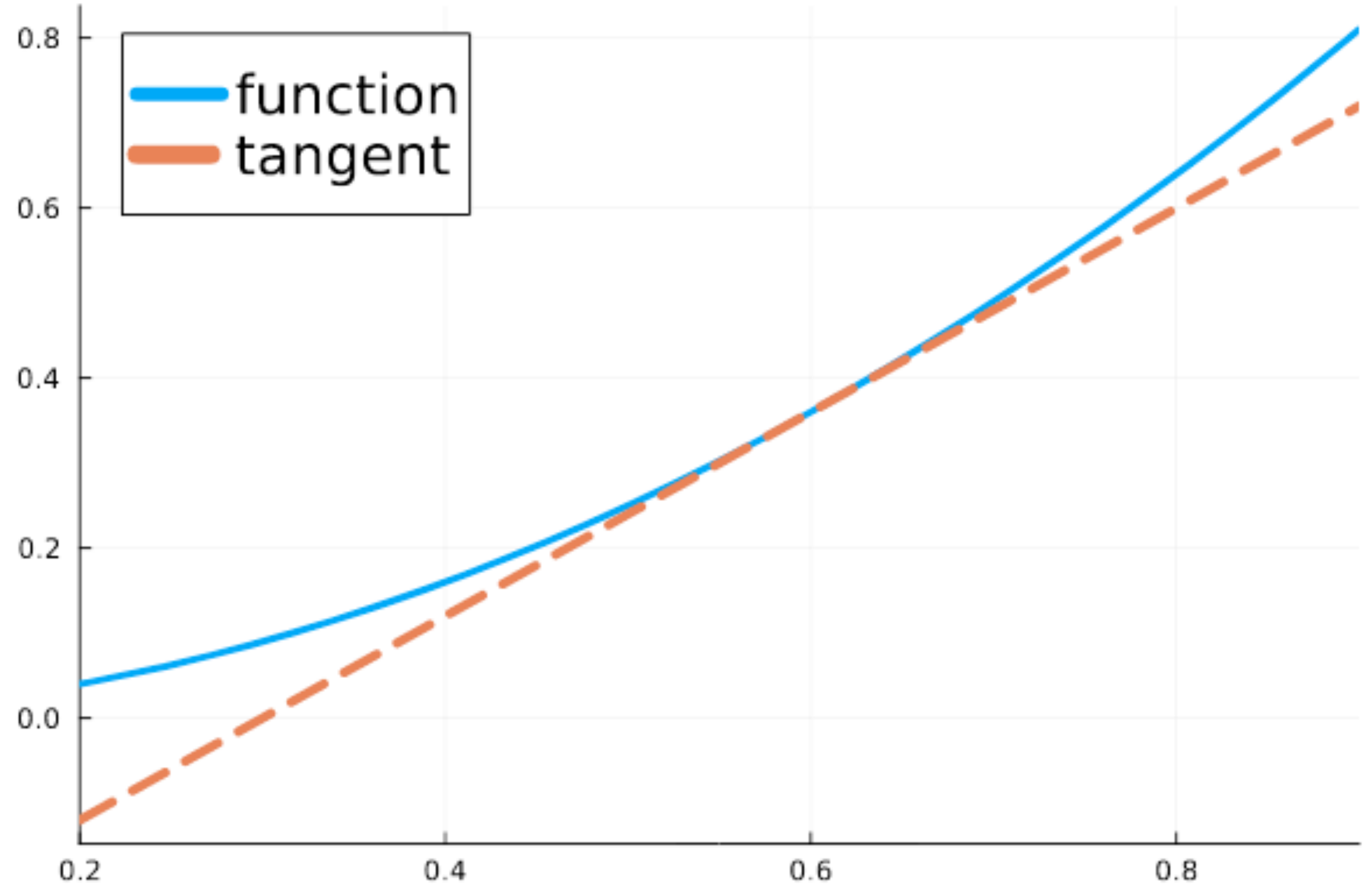
$$f'(x) = 0$$

at minimum…

Necessary to have zero derivative at minimum

If there is a slope, you can
go down the slope…

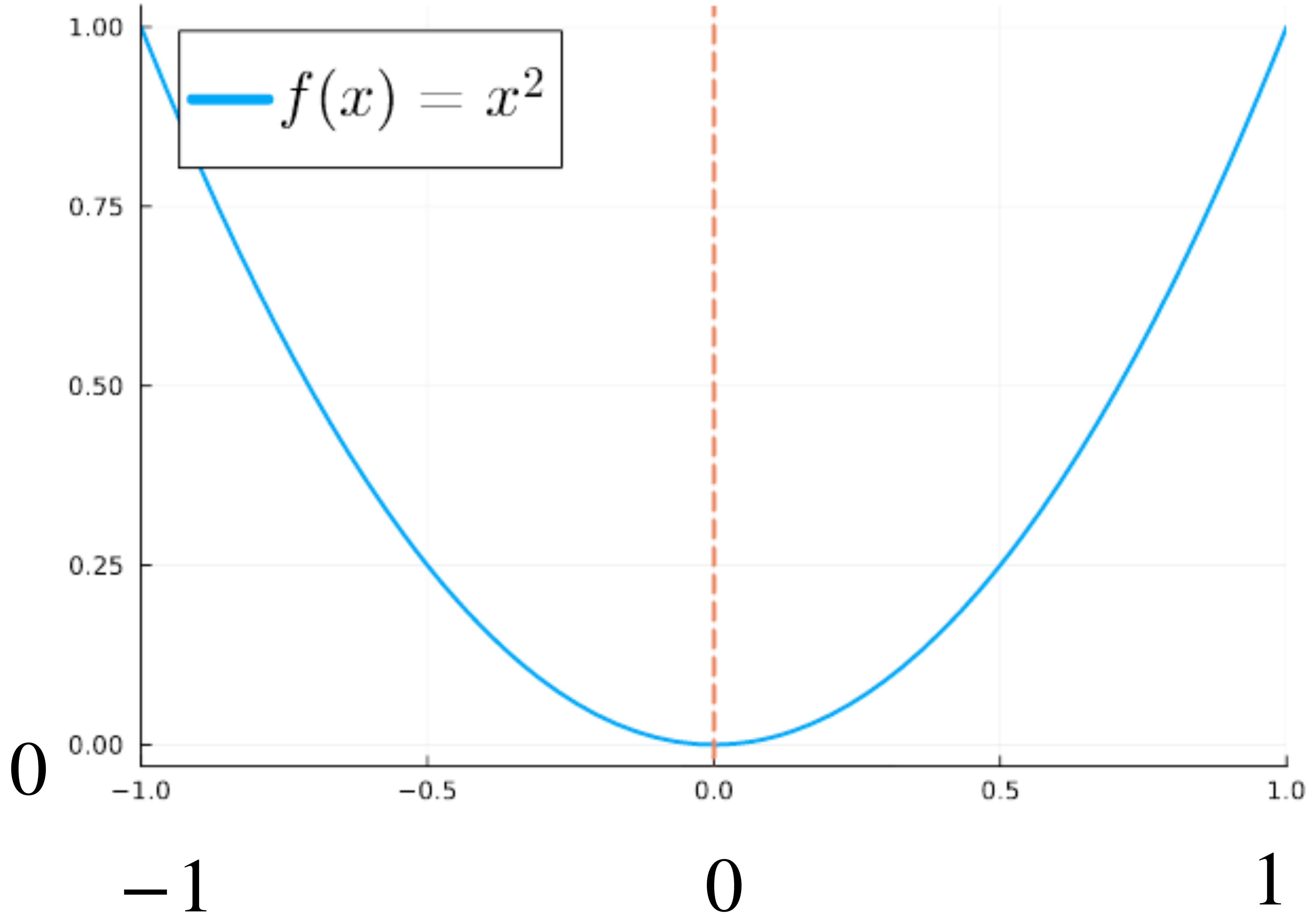
$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x$$



$f'(x) > 0 \Rightarrow$ Take $\Delta x < 0$

$f'(x) < 0 \Rightarrow$ Take $\Delta x > 0$

Unconstrained optimisation



$f'(x) = 0$ is necessary condition for minimality

Is it sufficient?

No...

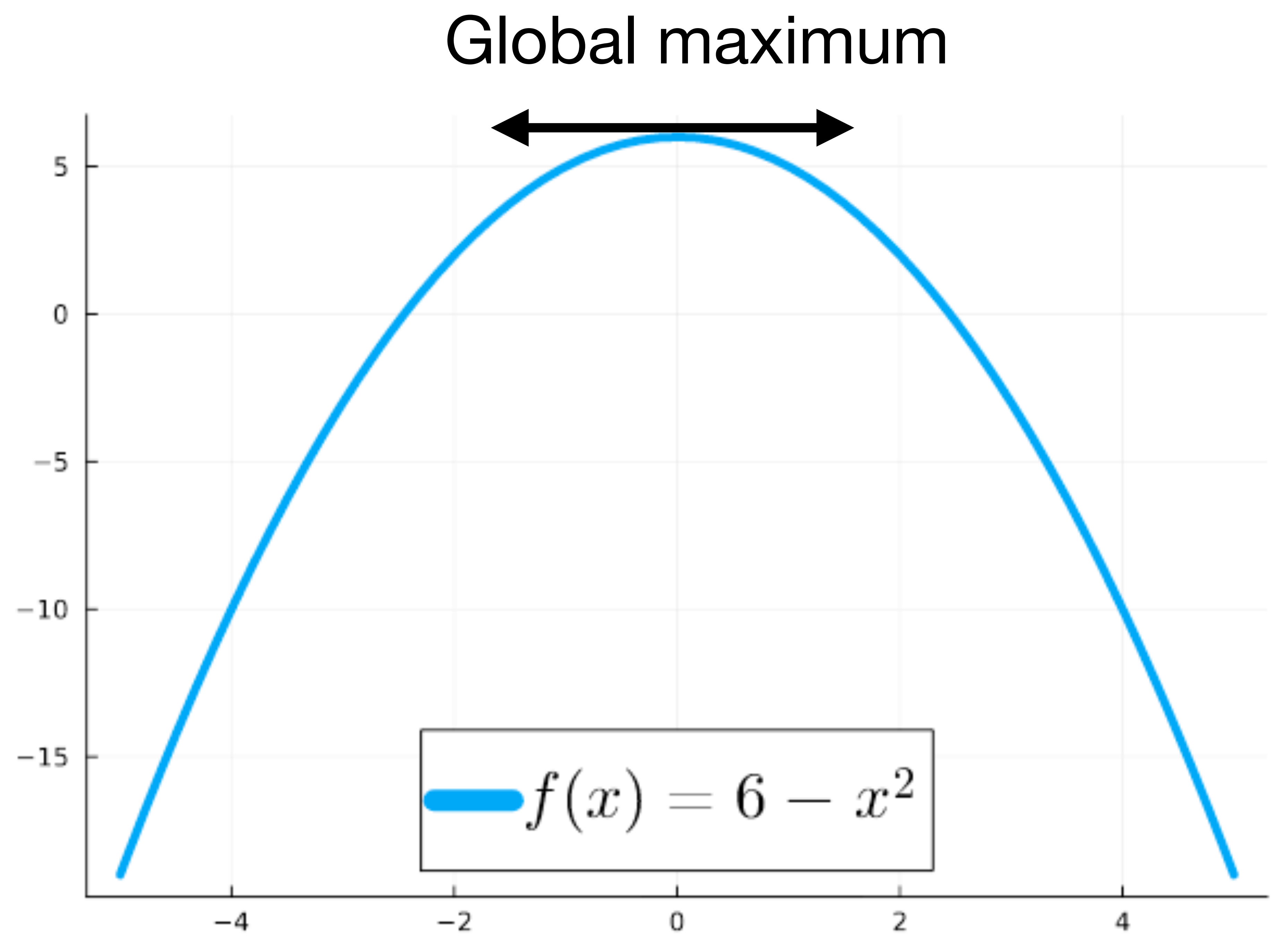
Maxima also have zero derivative

How do we know we're not at a maximum?

$f'(x)$: slope at x

$f''(x)$: remember what?

Curvature at x!



Sign of $f''(x)$?

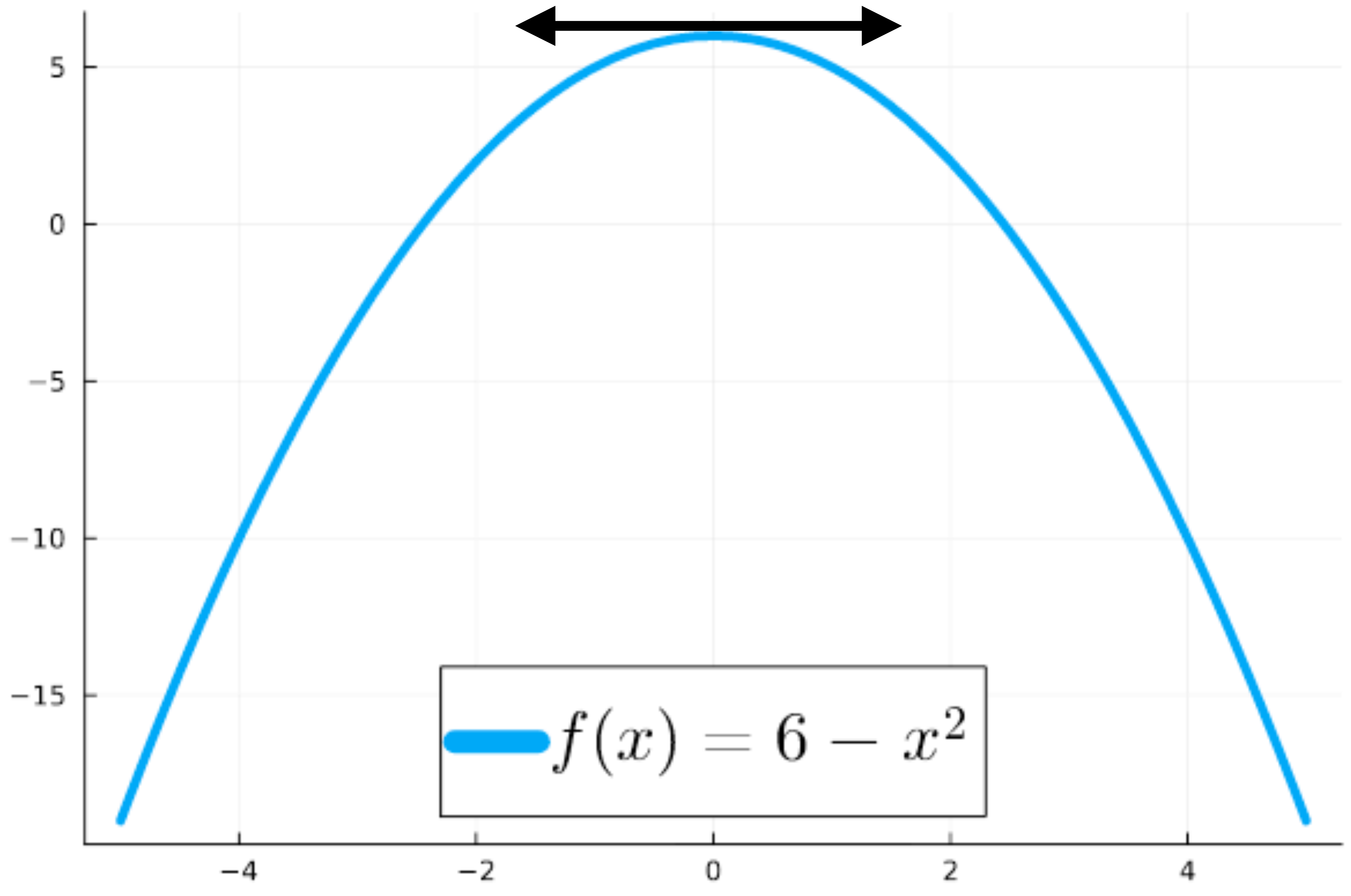
$f''(x) < 0$

Maxima also have zero derivative

Second derivative test

Maximum	Minimum
$f''(x) < 0$	$f''(x) > 0$

Global maximum



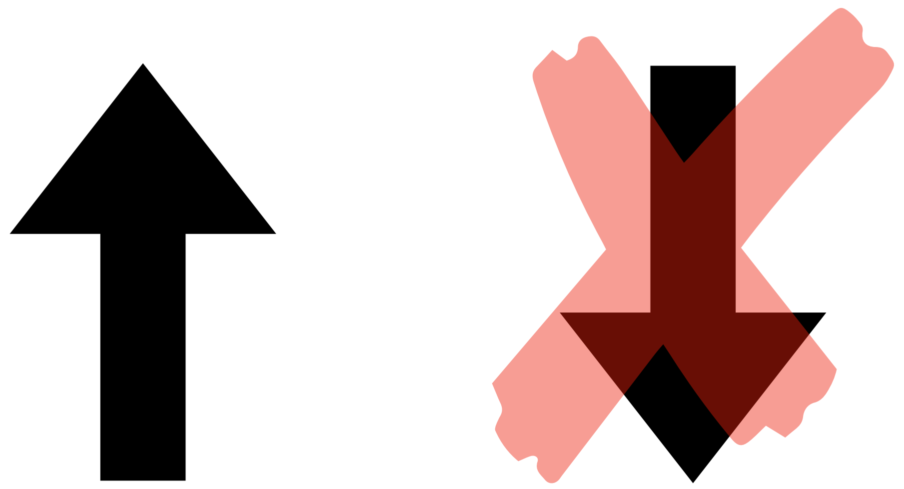
Sign of $f''(x)$?

$f''(x) < 0$

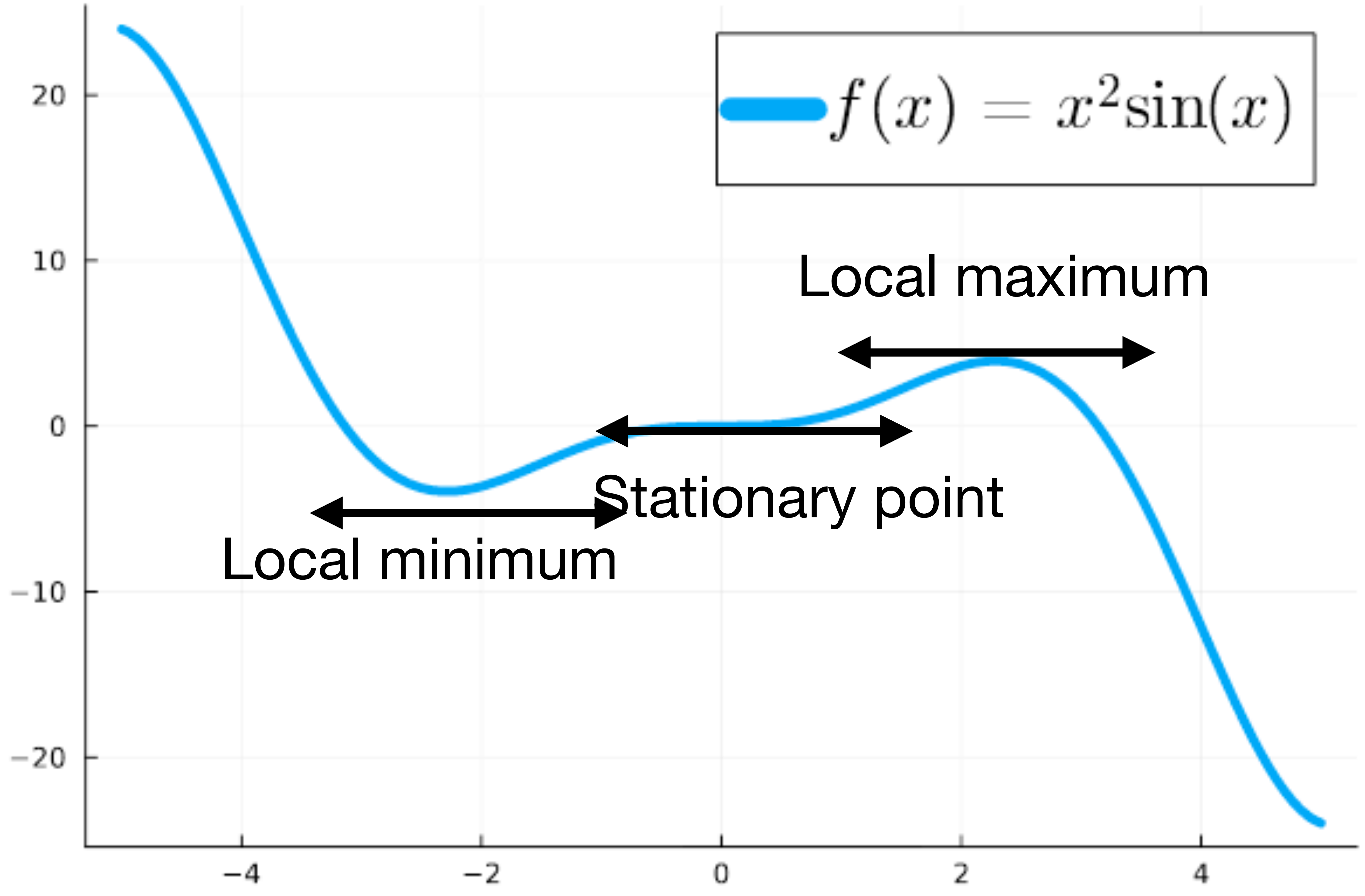
Not sufficient to have zero derivative at extremum

Zero derivative

"Stationary point"

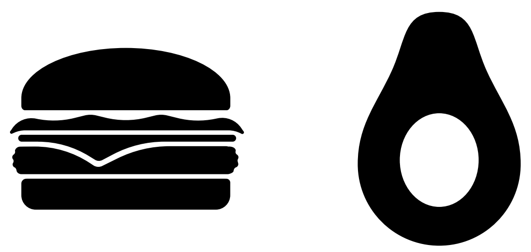


Extremum



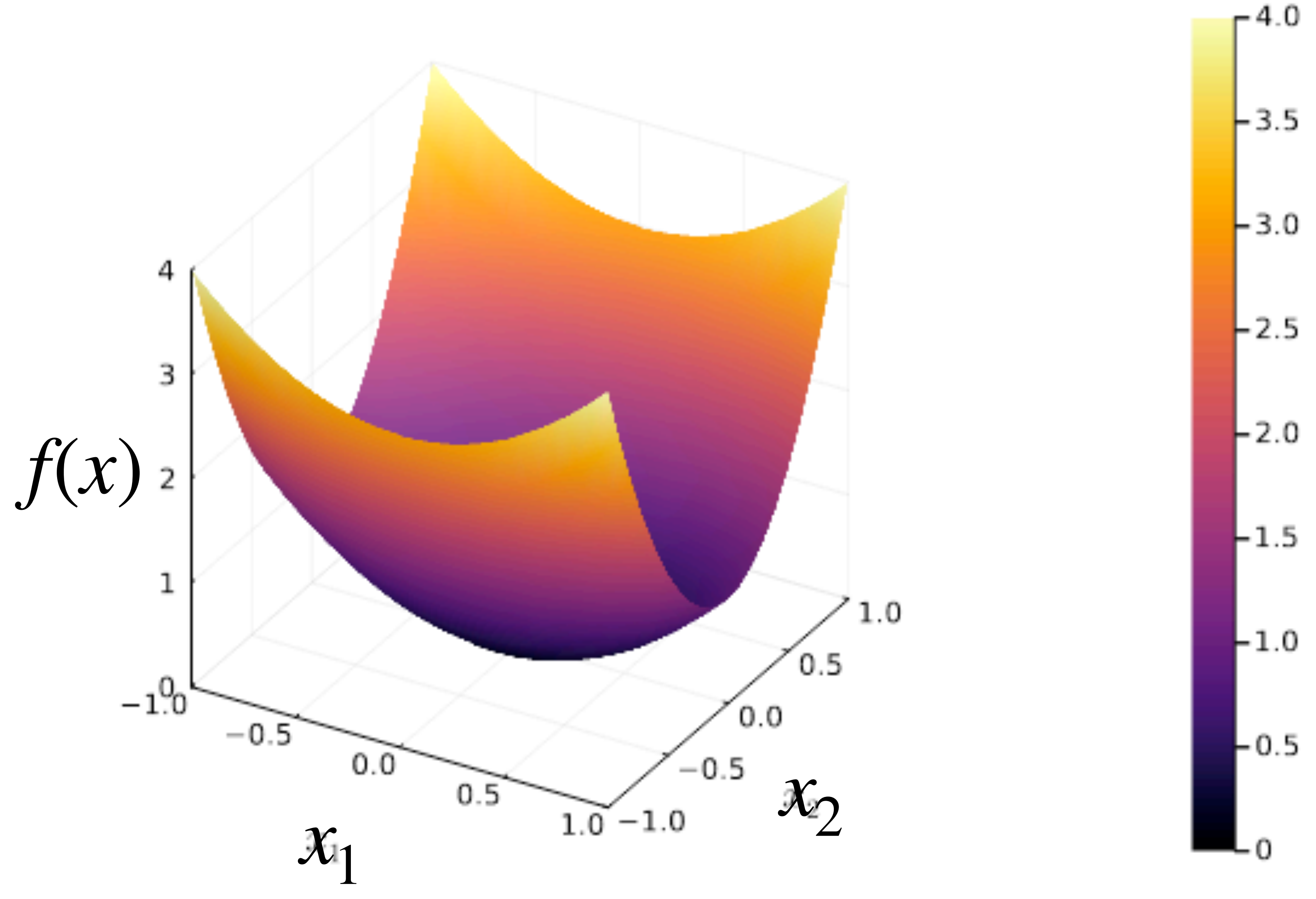
Multiple optimisation variables?

$$f(x) = x_1^2 + 3x_2^2 \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

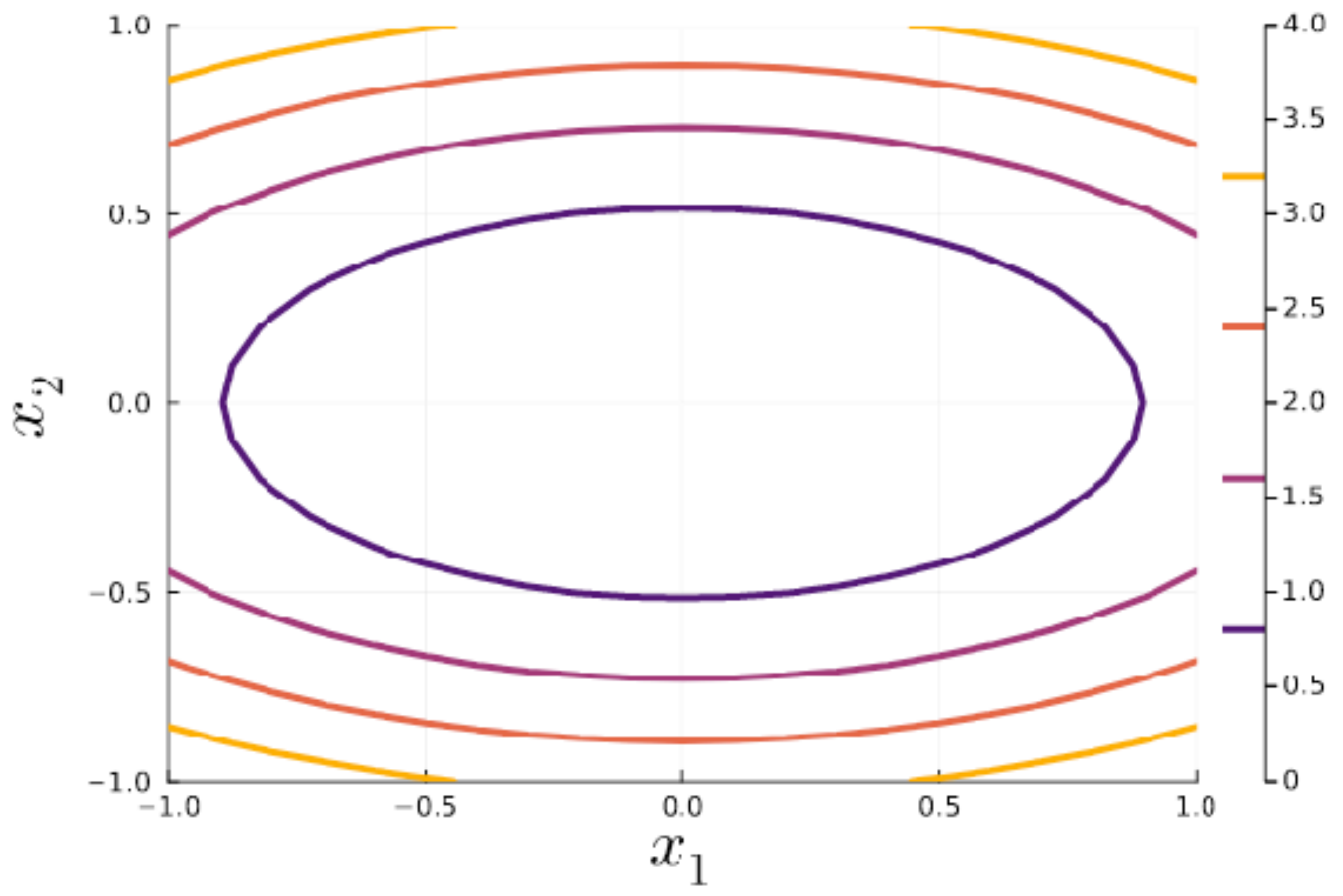


Clearly flat at minimum

Re-express using calculus?



Contour plot
("bird's eye view)



Multiple optimisation variables?

Subsequent lecture!

Need to figure out calculus with
multiple independent variables

Multivariable functions classification

Scalar function
(single output)

$$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$$

Vector-valued
function

$$f(x_1, x_2) = \begin{bmatrix} 3x_1^2 + 5x_1x_2 \\ x_2^2 + 3x_1\sin(x_2) \end{bmatrix}$$

Equivalent to

$$f(x, y) = 3x^2 + 5xy$$

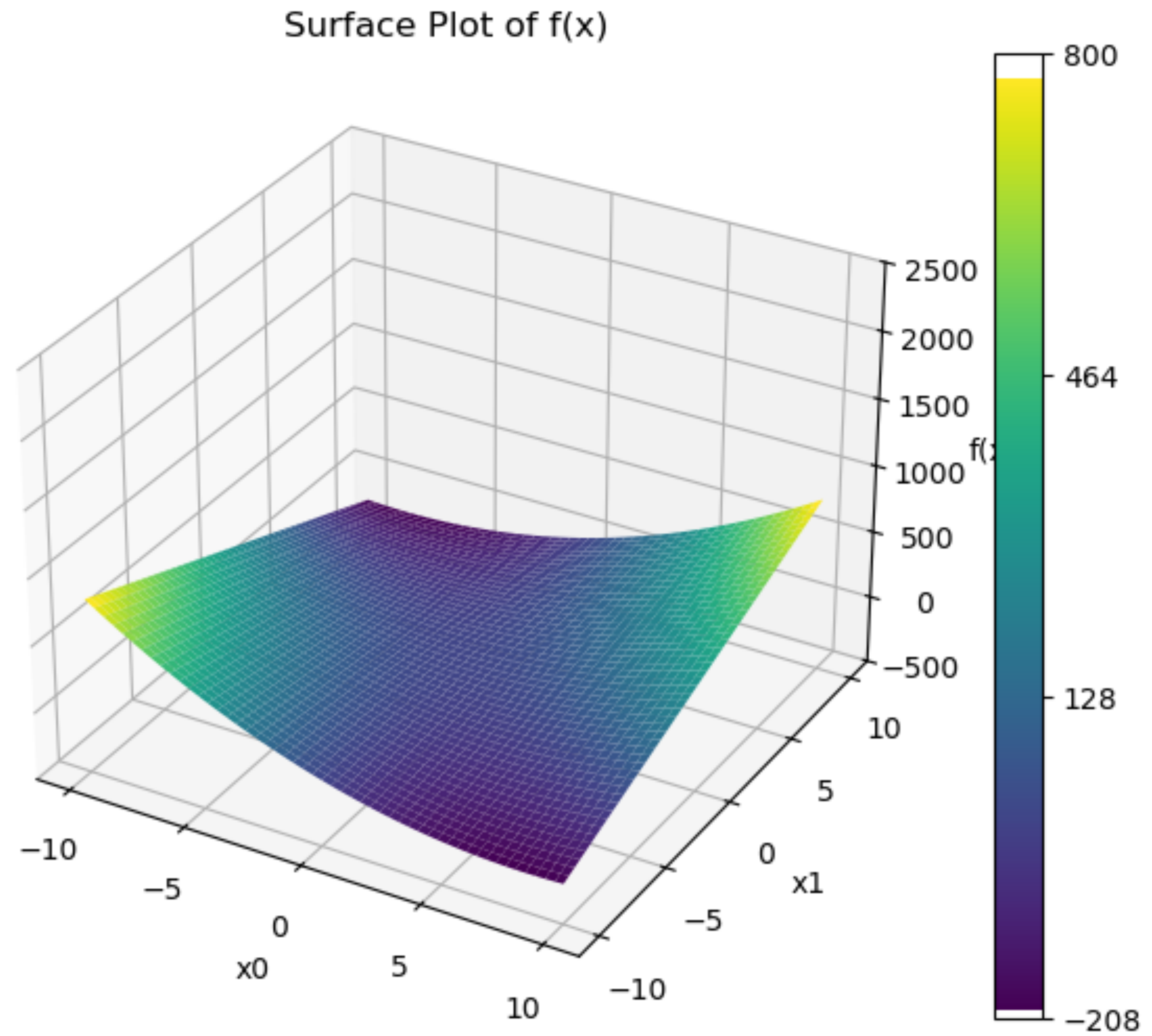
Matrix, tensor-
valued function…

Visualising scalar functions

Scalar function
(single output)

$$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$$

Surface plot: bad!

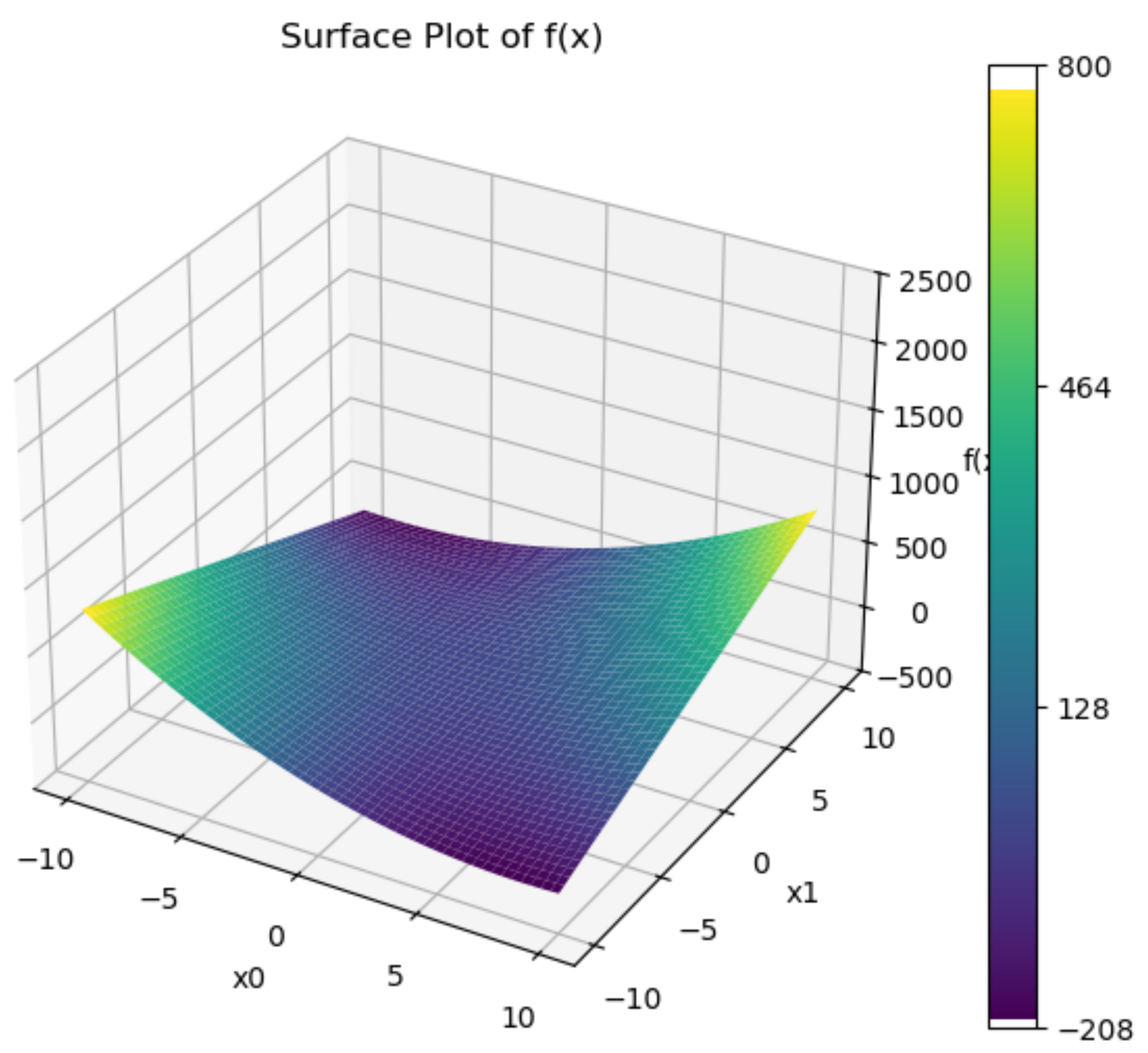


Color bar

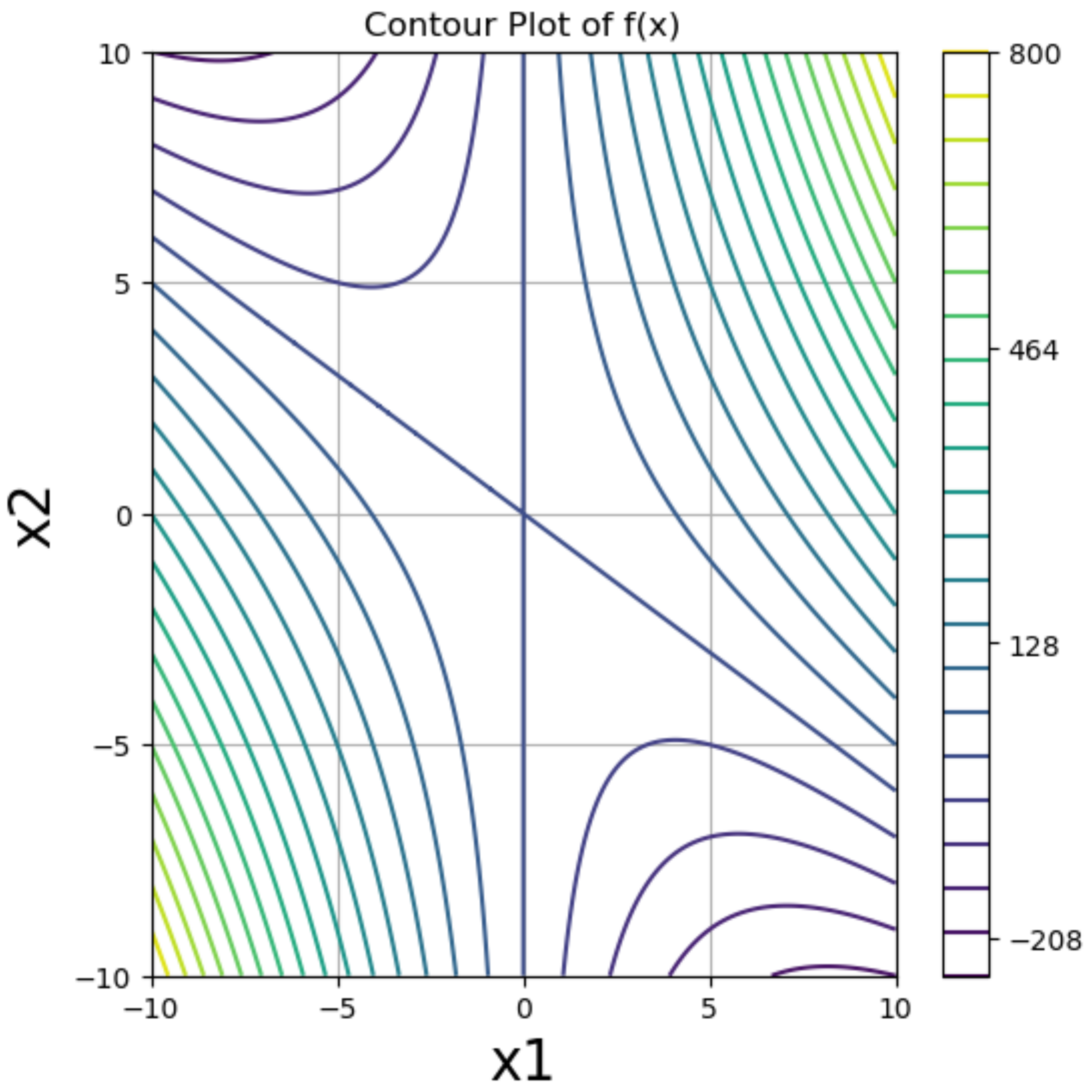
Visualising scalar functions

Scalar function
(single output)

$$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$$



Contour plot: better!



Looking at surface from above

What do lines represent?

Level sets:
$f(x) = c$

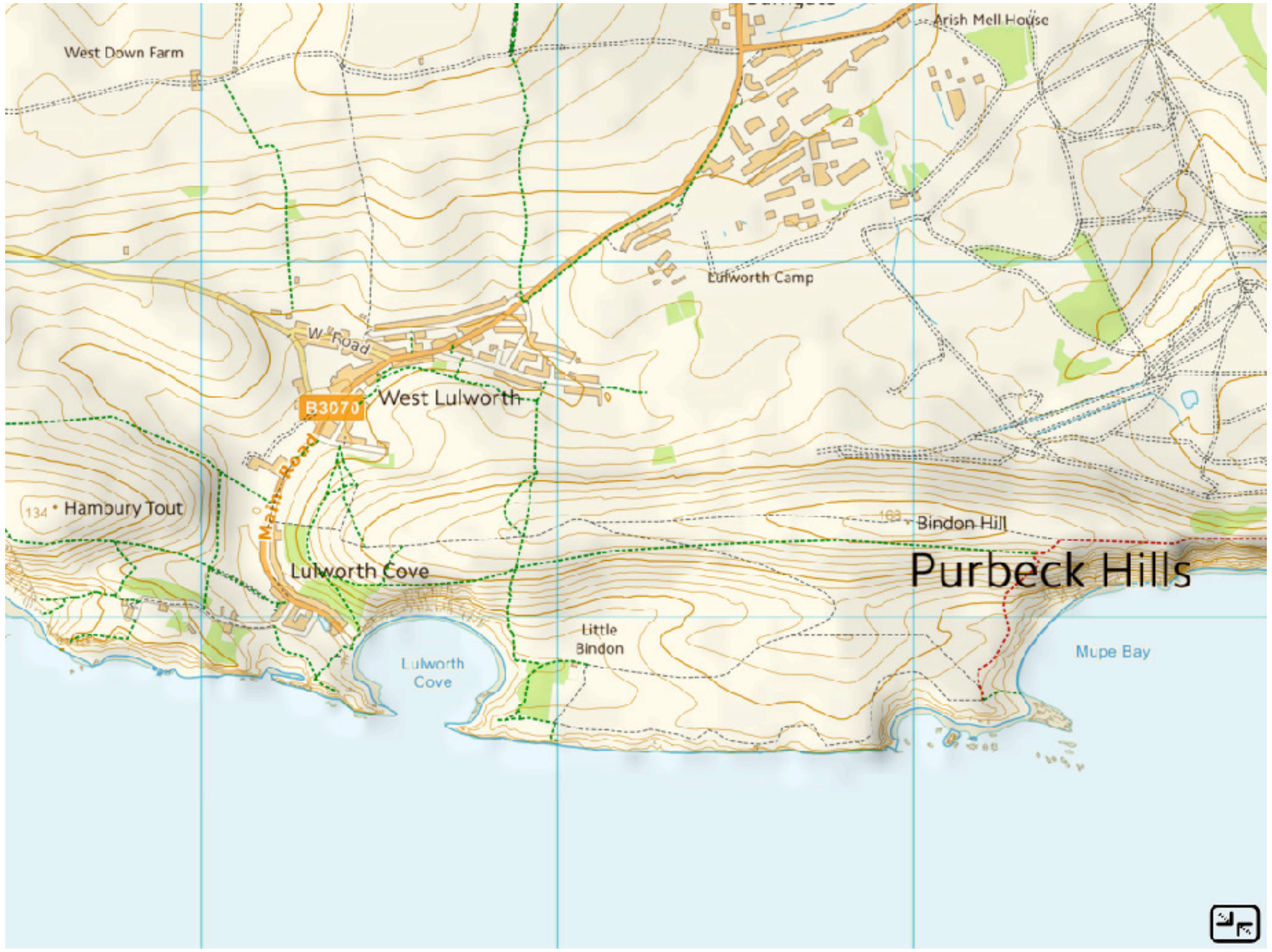
Common contour map

$z = f(x_1, x_2)$

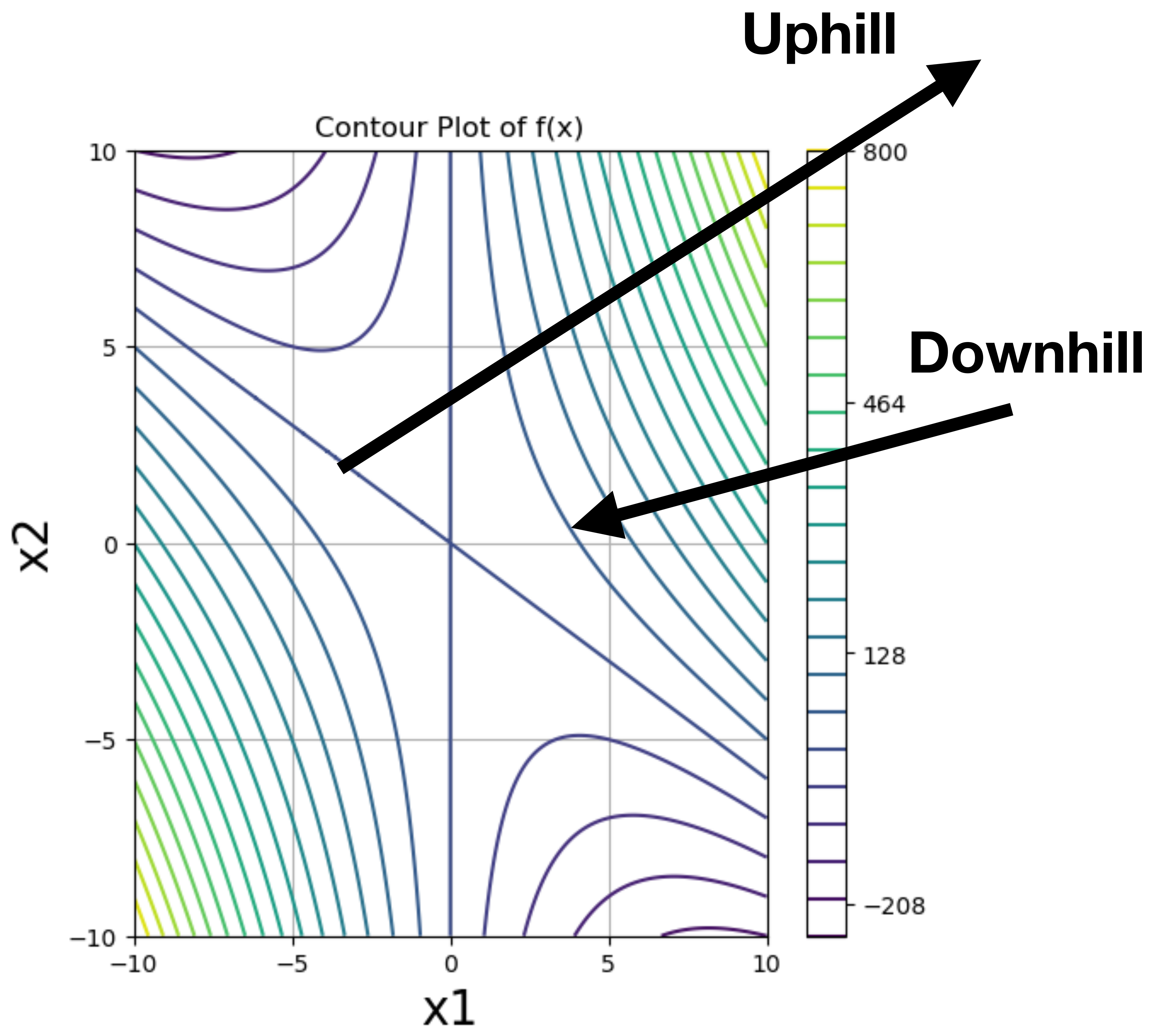
z : altitude

x_1 : longitude

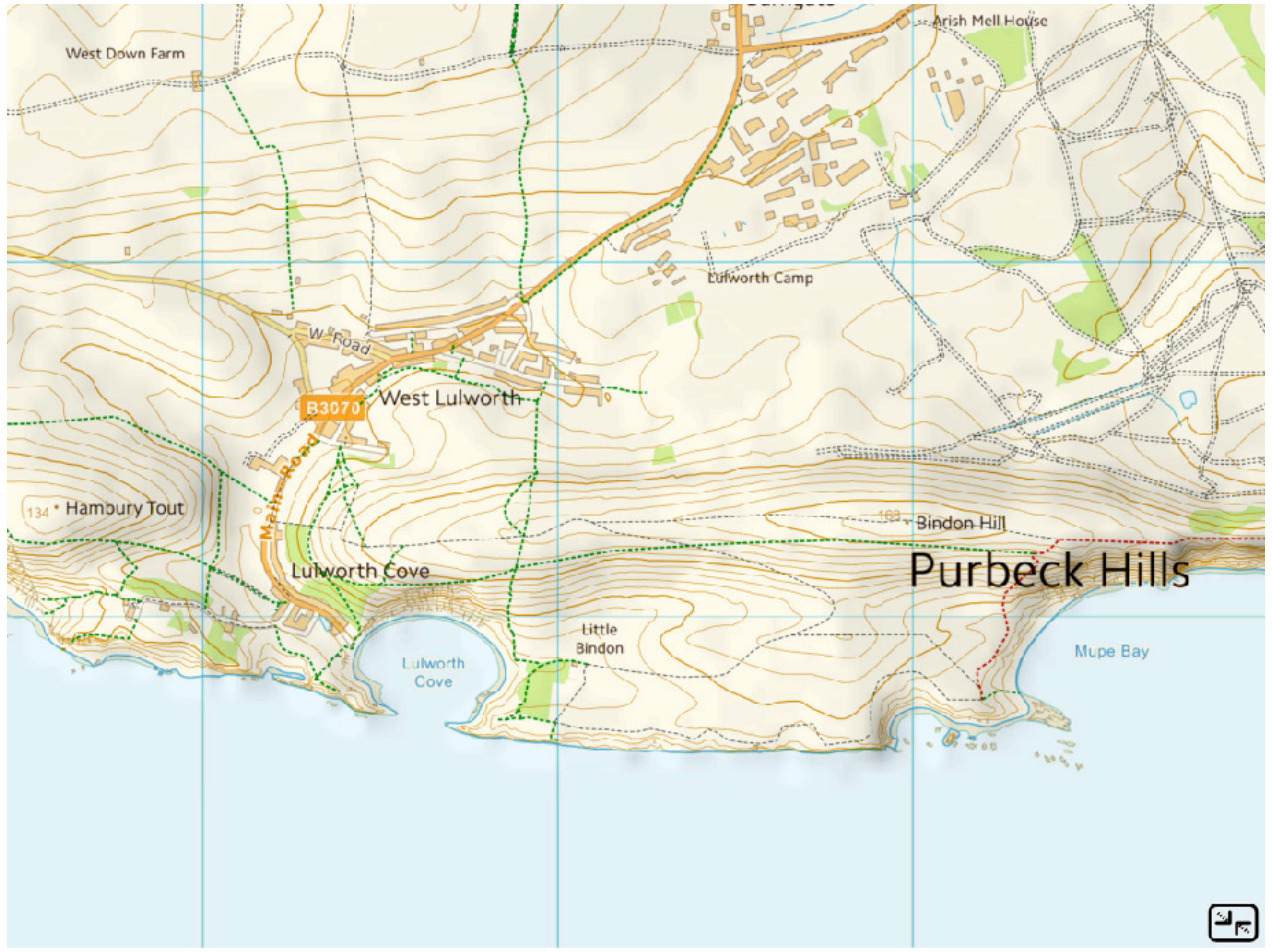
x_1 : latitude



Derivatives are now steepness
in a direction



How do we incorporate direction in our notion of derivatives?



Cross sections

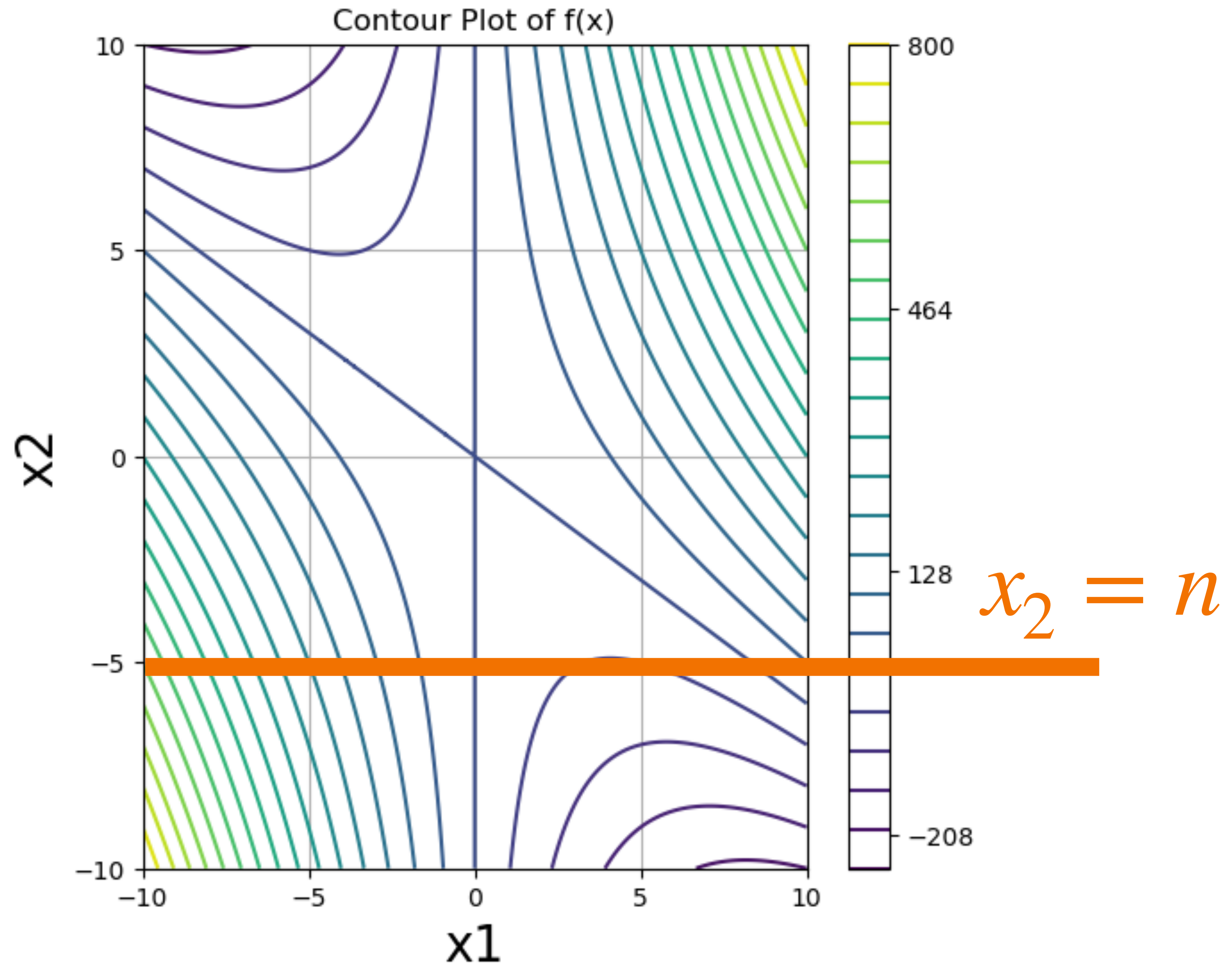
Scalar function
(single output)

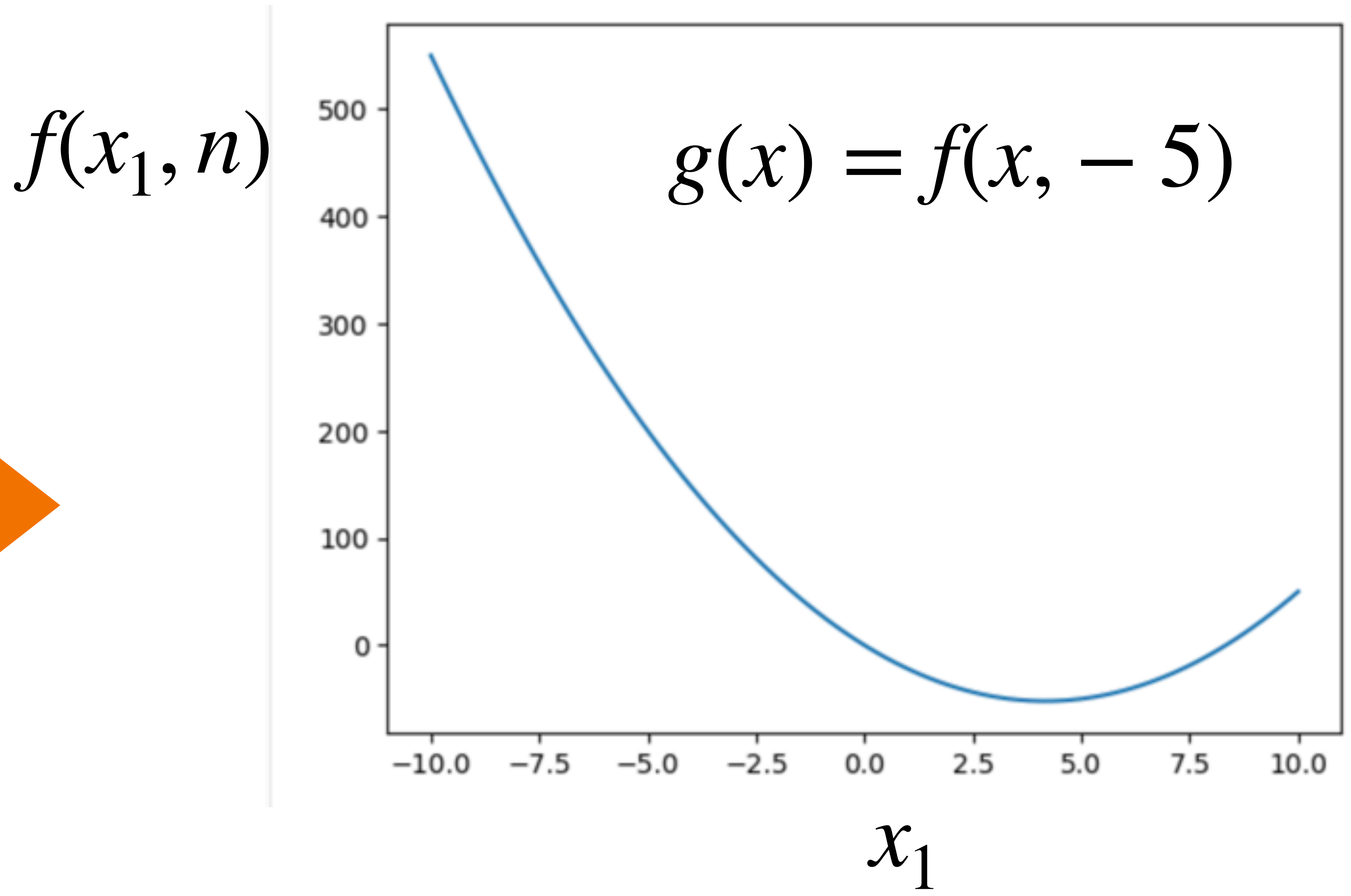
$$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$$

```python
def f(x):
    return 3*x[0]**2 + 5*x[0]*x[1]

def g(n):
    def newf(x):
        return f([x, n])
    return newf
g(-5)(3)
```

$$= -48$$





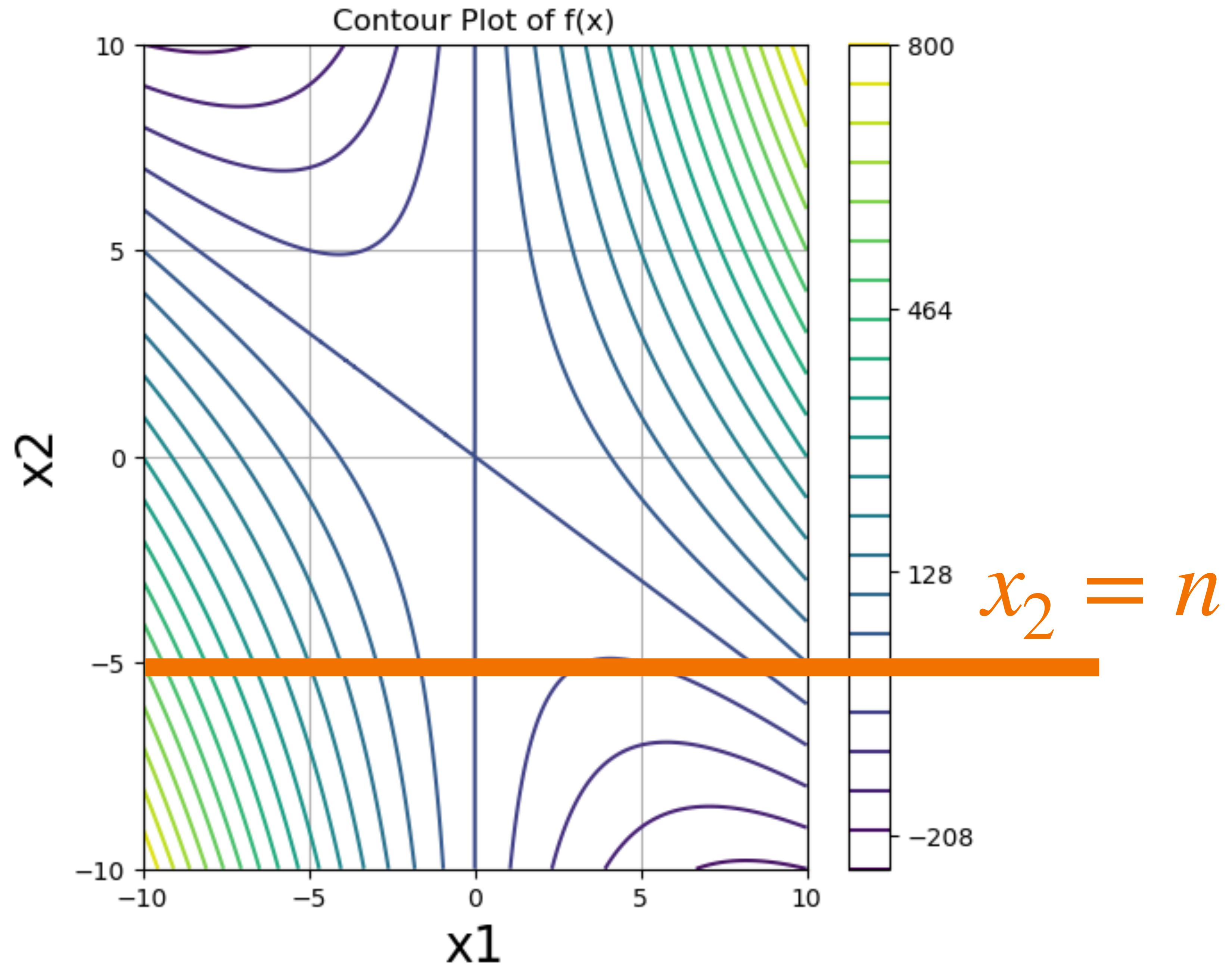
Derivatives are possible if…

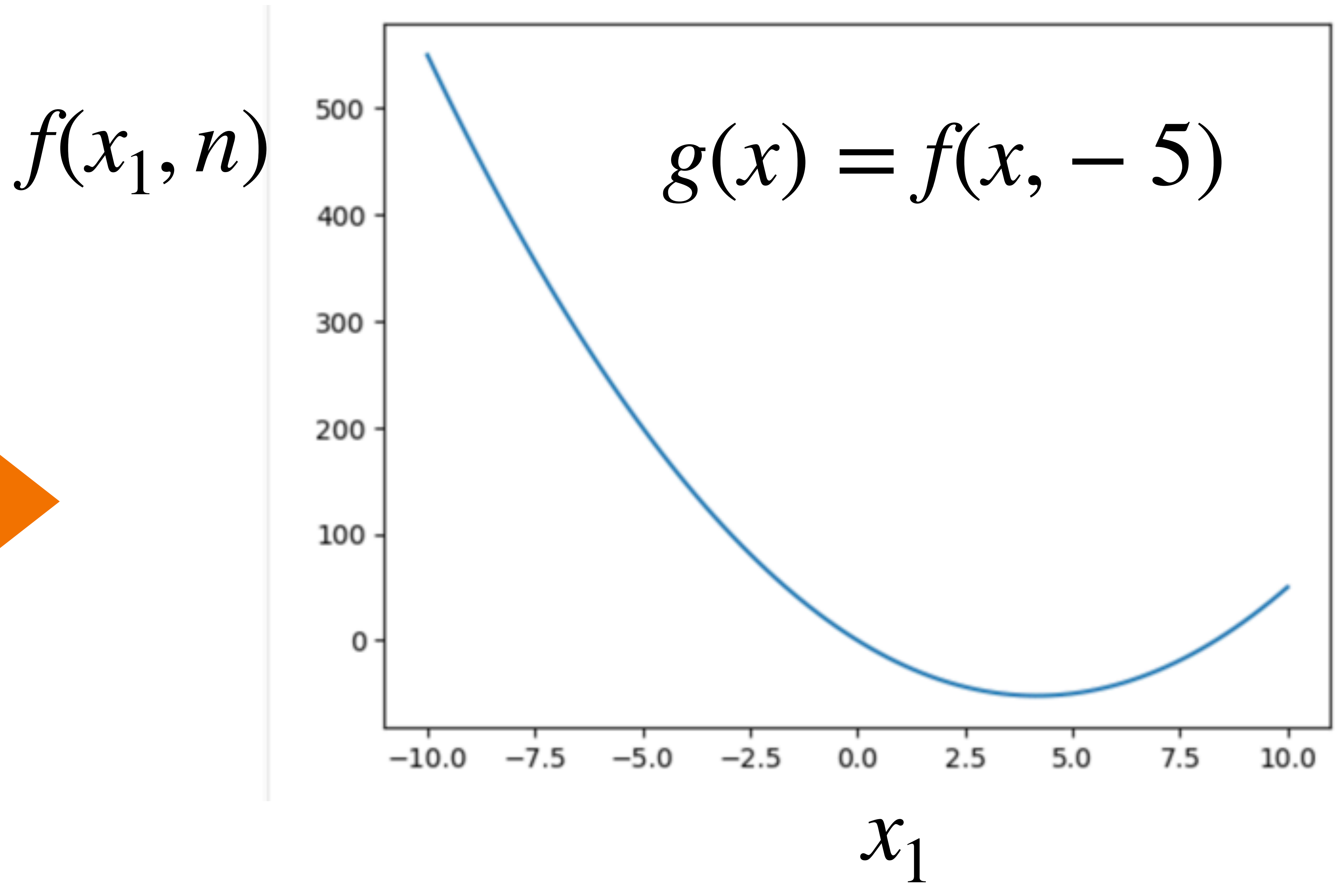
Scalar function
(single output)

$$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$$

Treat x_2 as a constant (number)

$$g(x) = f(x, -5) \qquad g'(x) = 6x^2 - 25$$



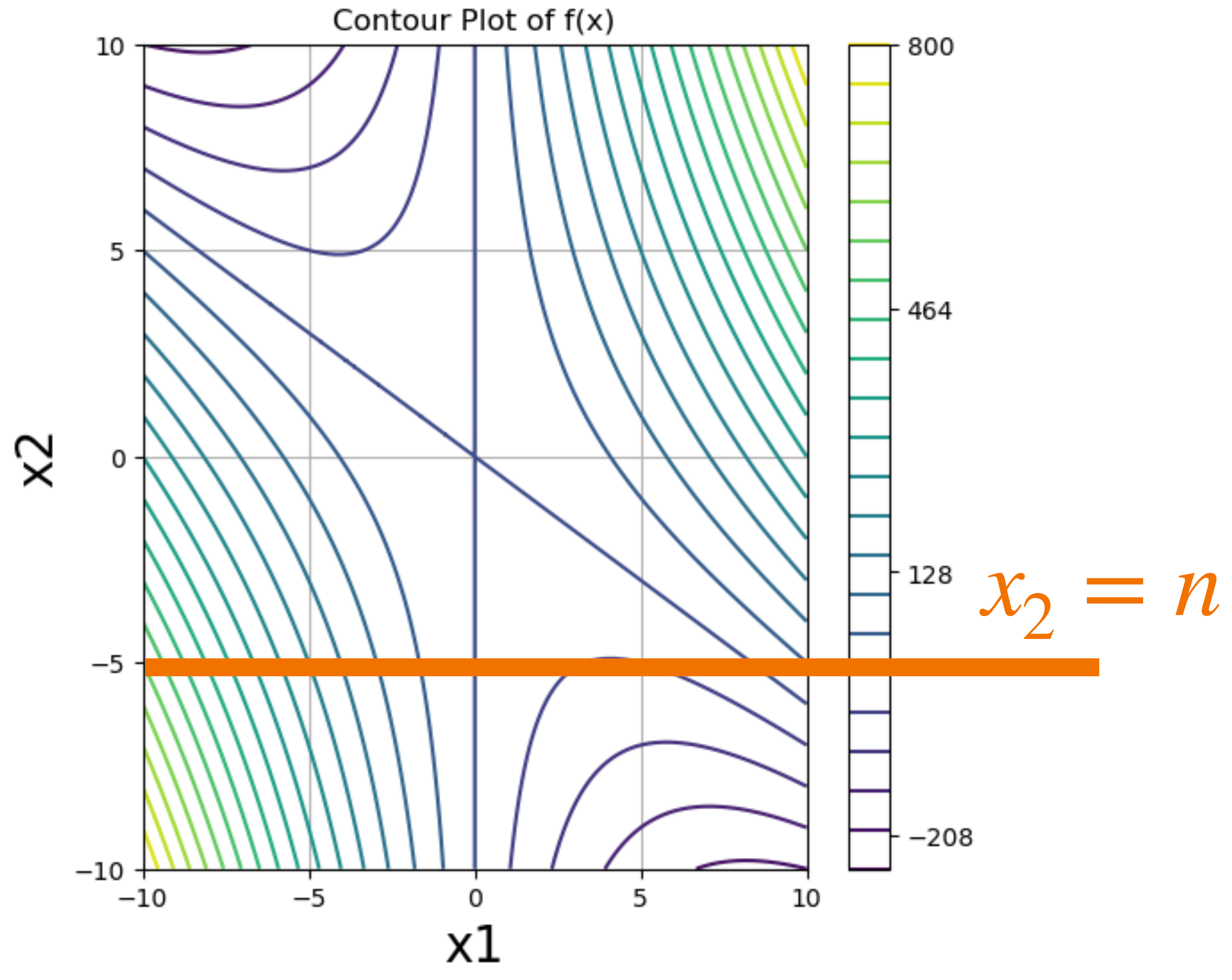


Partial derivatives

\frac{\partial f}{\partial x}

Scalar function
(single output)

$$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$$



What's the steepness if I only move in the x_1 direction?

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 6x_1 + 5x_2$$

Treat all independent variables that aren't x_1 as constant

Notice the different ∂ (partial / del)

Partial derivatives

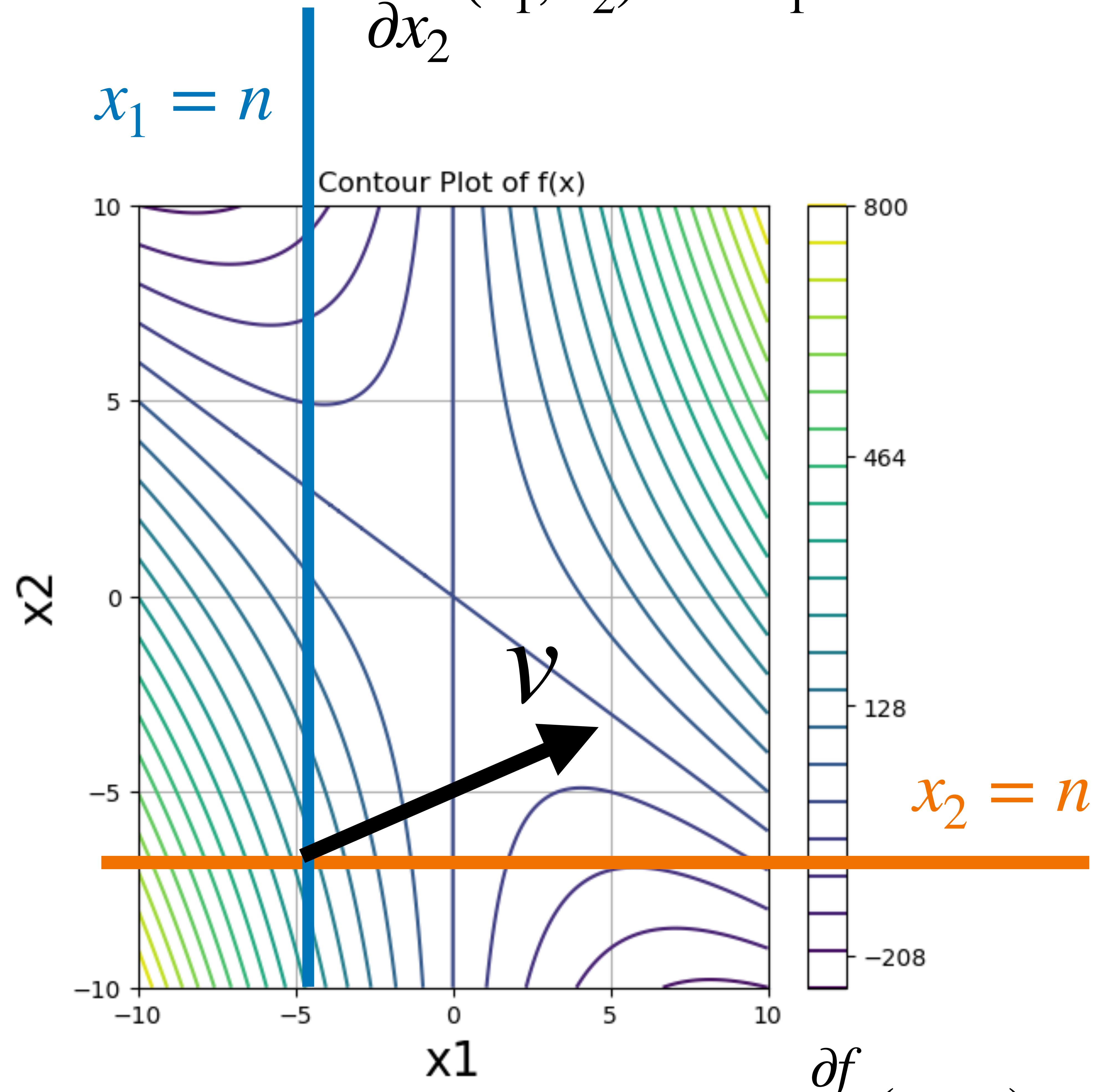
Scalar function
(single output)

$f(x_1, x_2) = 3x_1^2 + 5x_1x_2$

What about other directions?

e.g. $v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ $\qquad 3\frac{\partial f}{\partial x_1} + 1\frac{\partial f}{\partial x_2}$

$\frac{\partial f}{\partial x_2}(x_1, x_2) = 5x_1$



$\frac{\partial f}{\partial x_1}(x_1, x_2) = 6x_1 + 5x_2$

Why does this work

Curvature disappears if
you zoom in enough

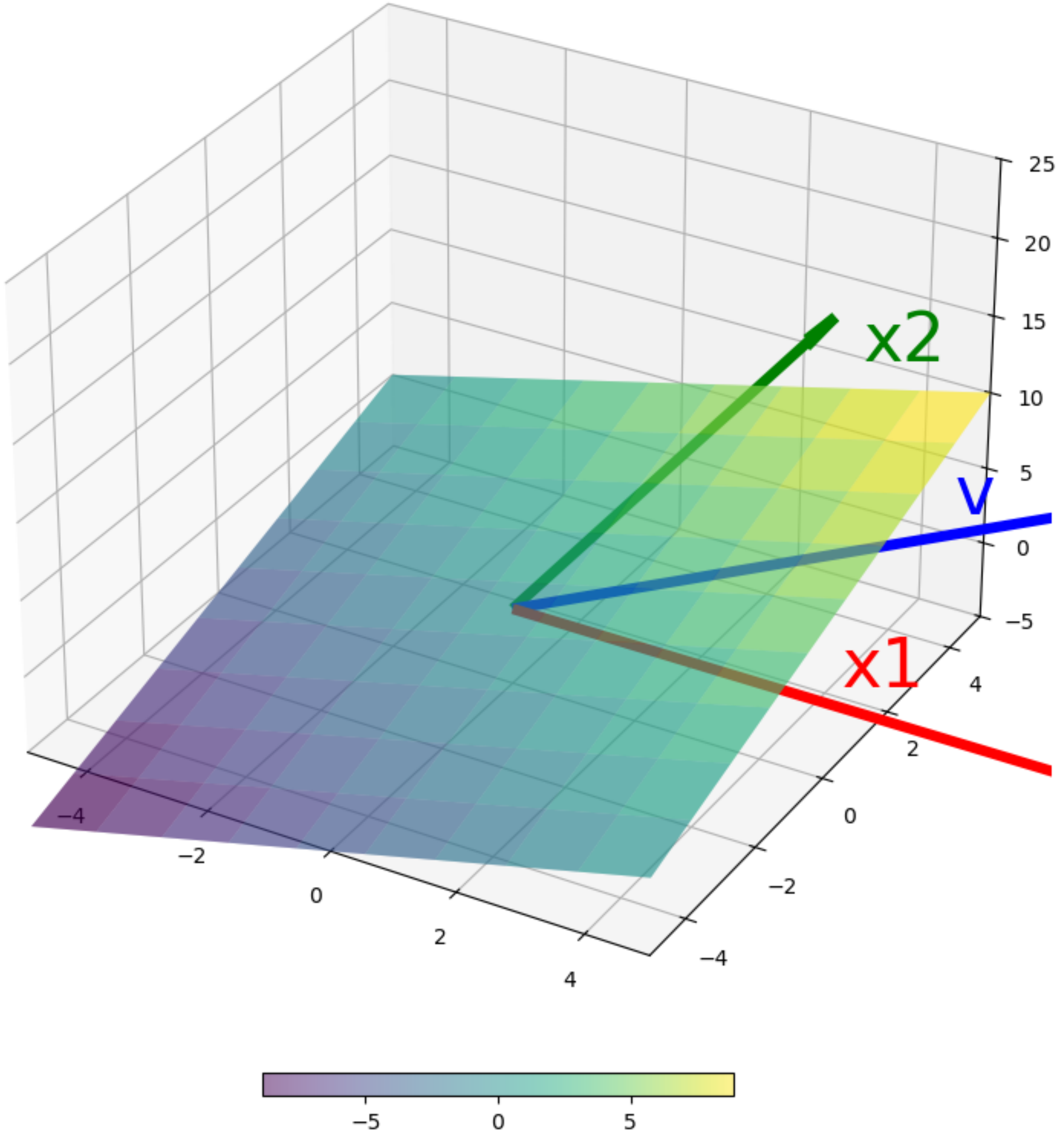
Steepness of a plane
is constant

"Walk 3 steps to the right, and 1 forward"

"Add steepnesses of each step"

e.g. $v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ $\qquad 3\frac{\partial f}{\partial x_1} + 1\frac{\partial f}{\partial x_2}$

Plane (no curvature)



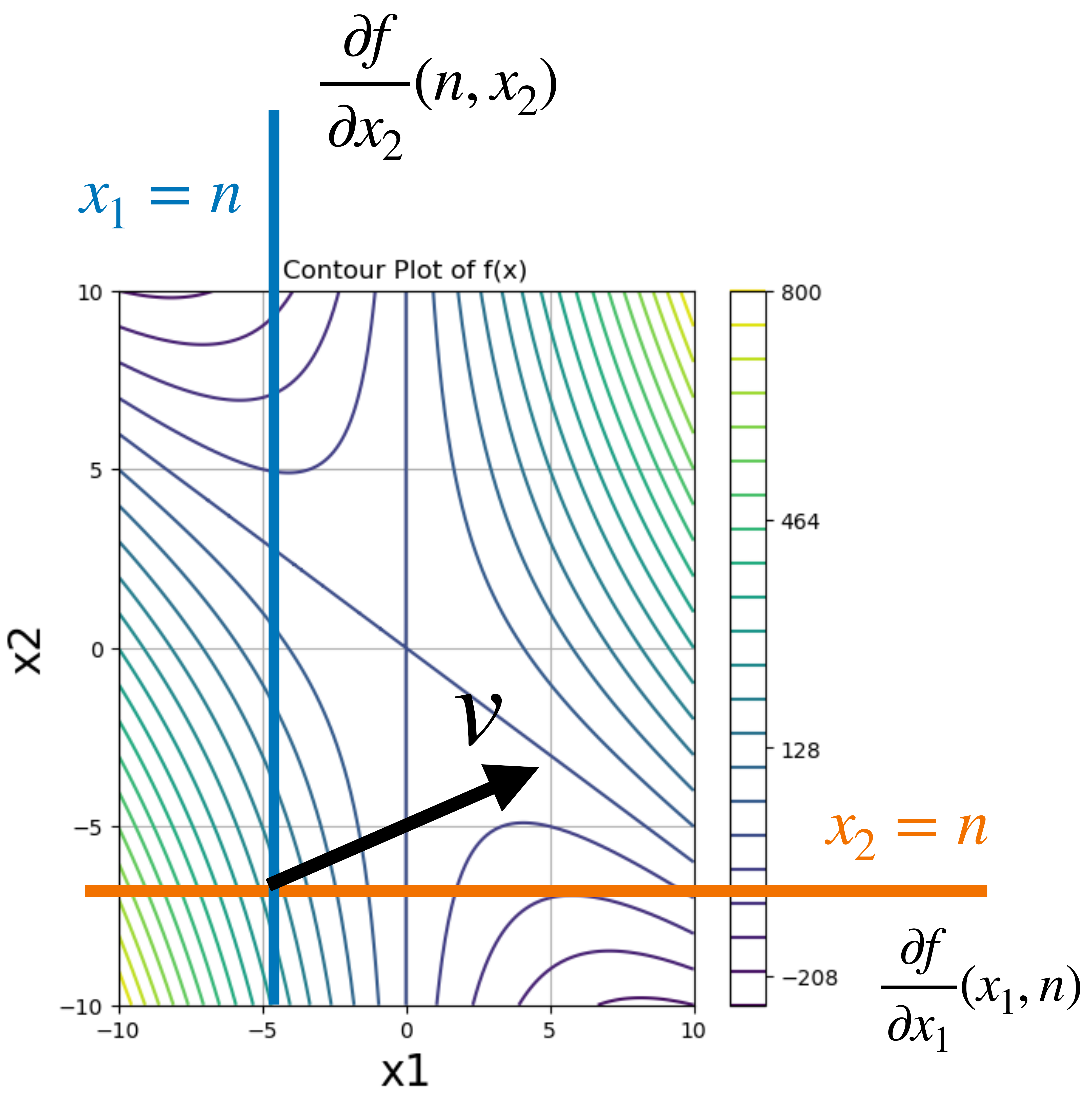
Directional derivatives

e.g. $v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ $\qquad 3\frac{\partial f}{\partial x_1} + 1\frac{\partial f}{\partial x_2}$

$$= [3 \quad 1] \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}$$

Dot product!

Vector of partial derivatives / gradient



Notational chaos

e.g. $v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ $\quad 3\frac{\partial f}{\partial x_1} + 1\frac{\partial f}{\partial x_2}$

$= [3 \quad 1]\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}$

Dot product!

Vector of partial derivatives / gradient

Pasted from wikipedia: directional derivatives

$$\nabla_{\mathbf{v}} f(\mathbf{x}) = f'_{\mathbf{v}}(\mathbf{x}) = D_{\mathbf{v}} f(\mathbf{x}) = Df(\mathbf{x})(\mathbf{v}) = \partial_{\mathbf{v}} f(\mathbf{x}) = \mathbf{v} \cdot \nabla f(\mathbf{x}) = \mathbf{v} \cdot \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}.$$

We will use: 'nabla'

\nabla

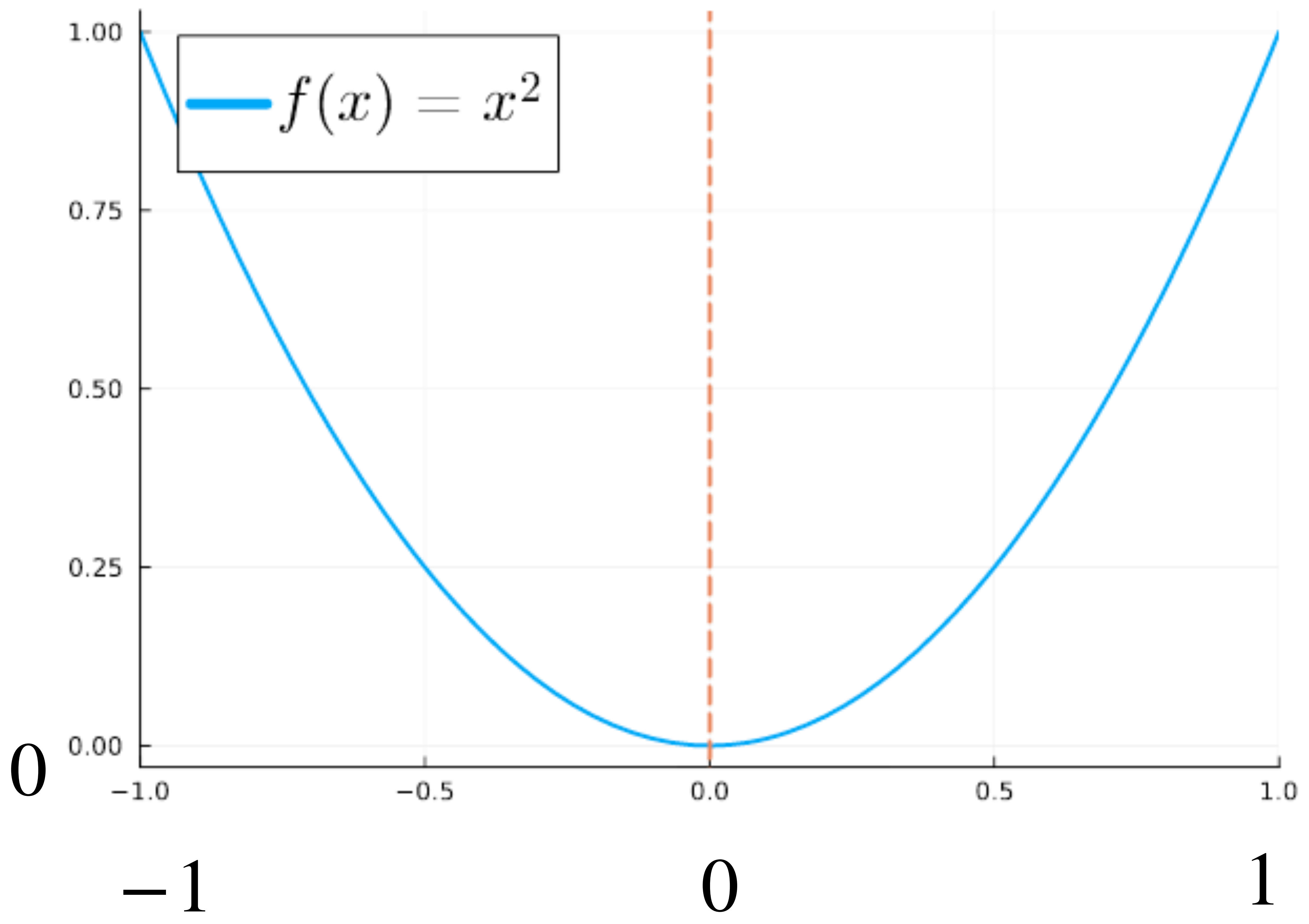
$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x, y) \\ \frac{\partial f}{\partial x_2}(x, y) \end{bmatrix}$$

How many inputs?

How many outputs?

$$\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$$

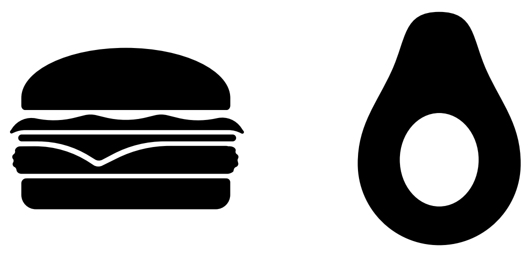
Unconstrained optimisation



$f'(x) = 0$ is necessary condition for minimality

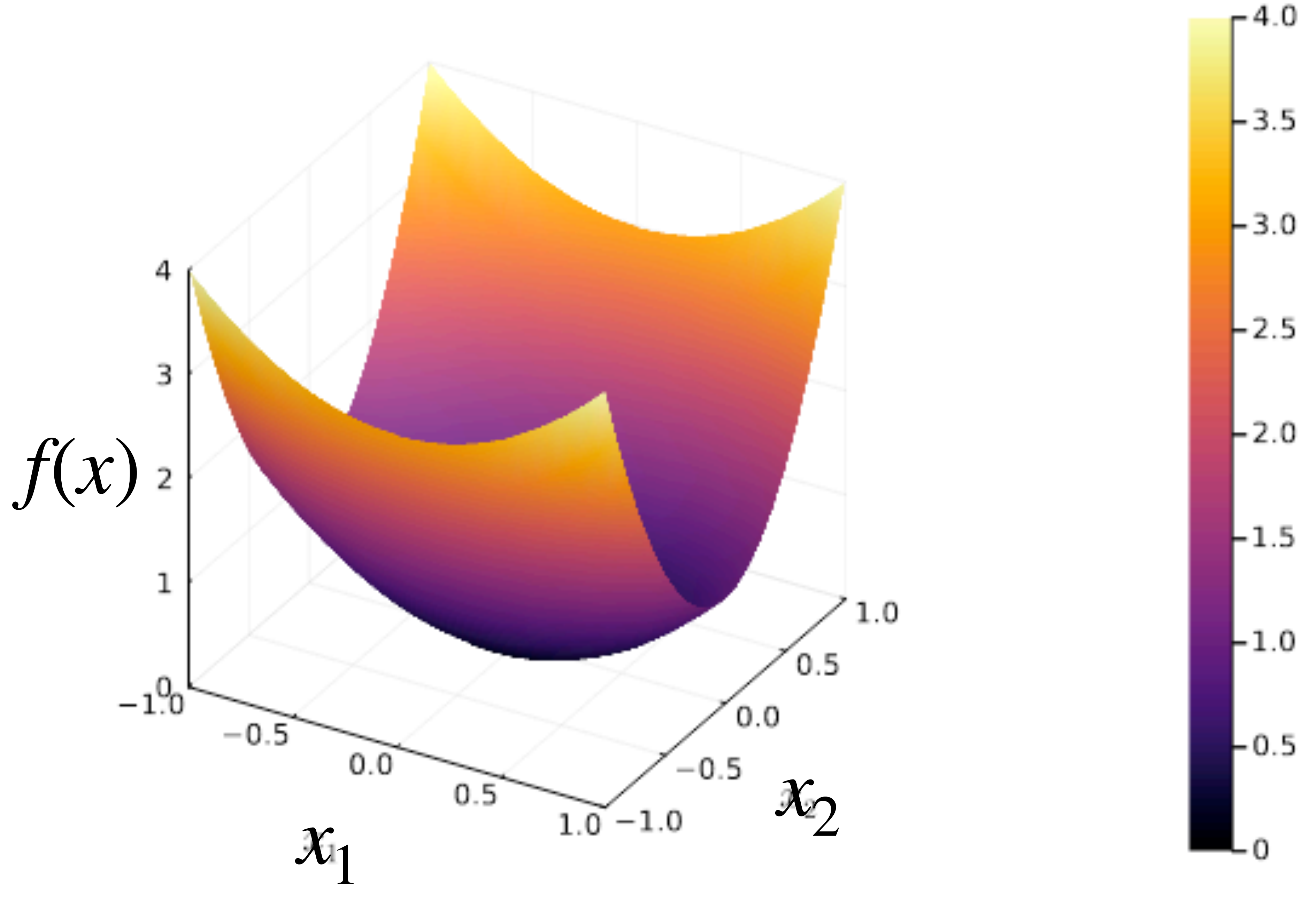
Multiple optimisation variables?

$$f(x) = x_1^2 + 3x_2^2 \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

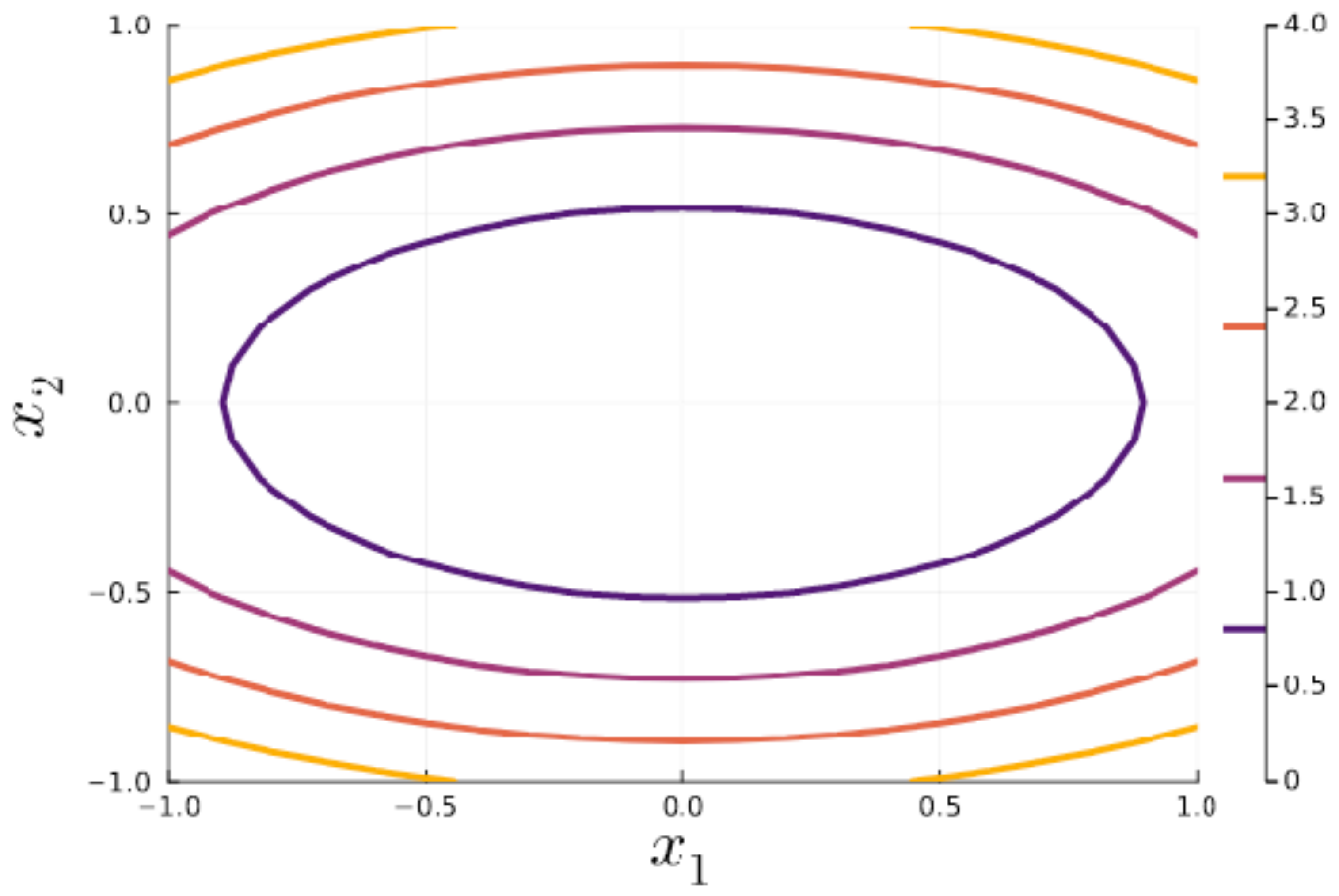


Necessary condition on minimum?

Directional derivative is zero in all directions

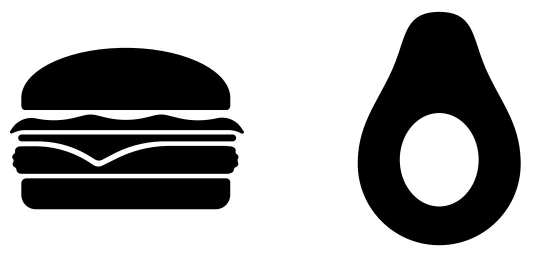


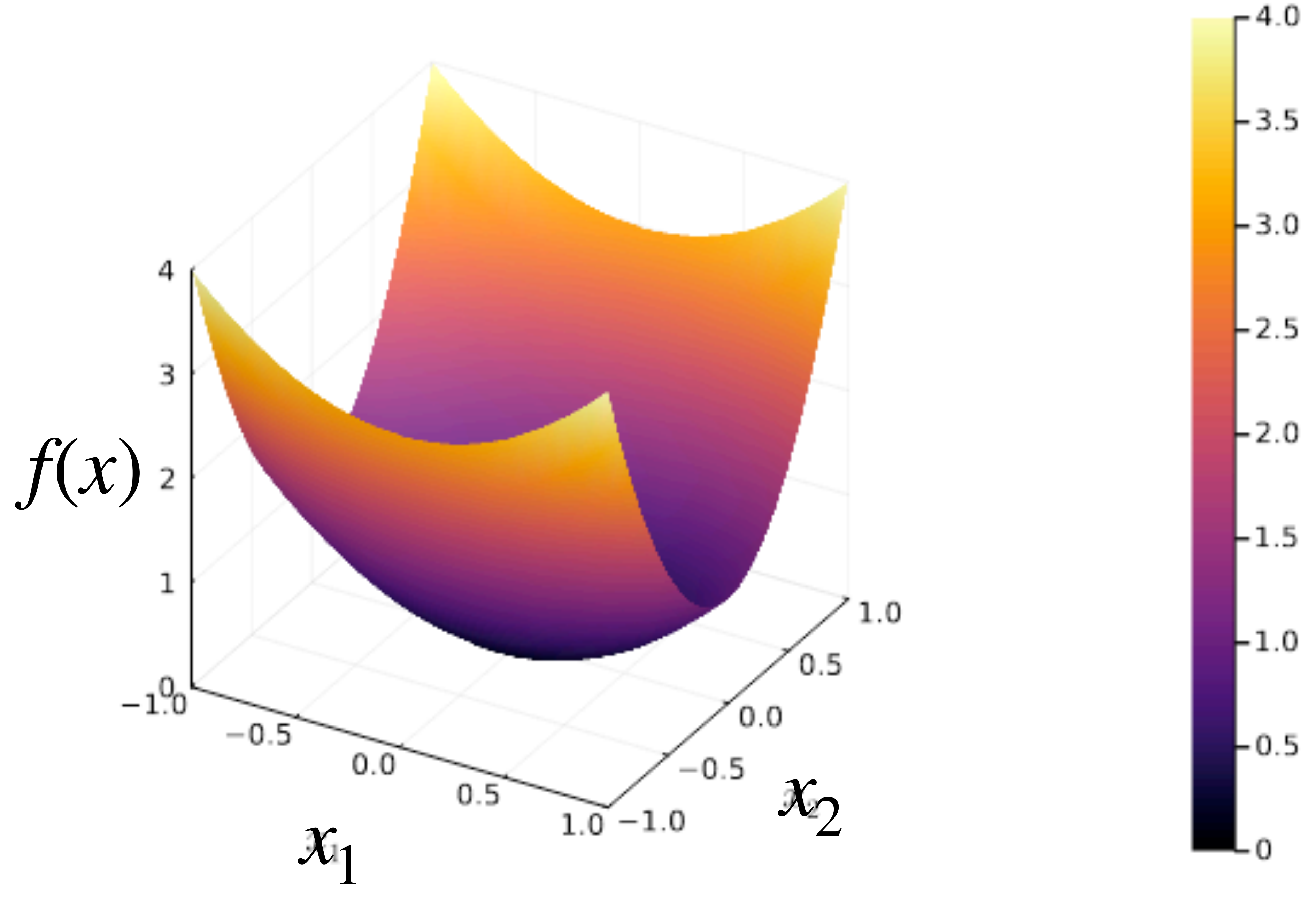
Contour plot
(“bird’s eye view)



Multiple optimisation variables?

$$f(x) = x_1^2 + 3x_2^2 \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$





Necessary condition on minimum?

Directional derivative is zero in all directions

$$\langle \nabla f, v \rangle = [v_1 \quad v_2] \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = 0 \quad \forall v$$

Gradient ∇f is zero!

Practice

Function	Derivative
$y = x^n$	$y' = nx^{n-1}$

common notation: x assumed to be vector

$$f(x) = 3x_1^2 + 4x_1x_2^2 + 2x_1 + 4x_3x_2$$

$$\frac{\partial f}{\partial x_1} = \quad 6x_1 + 4x_2^2 + 2$$

$$\frac{\partial f}{\partial x_2} = \quad 8x_1x_2 + 4x_3$$

$$\frac{\partial f}{\partial x_3} = \quad 4x_2$$

$$\nabla f(x, y) = 0\,?$$

$$x_2 = \frac{0}{4} = 0$$

$$6x_1 = -2 : \quad x_1 = -\frac{1}{3}$$

$$4x_3 = 0$$

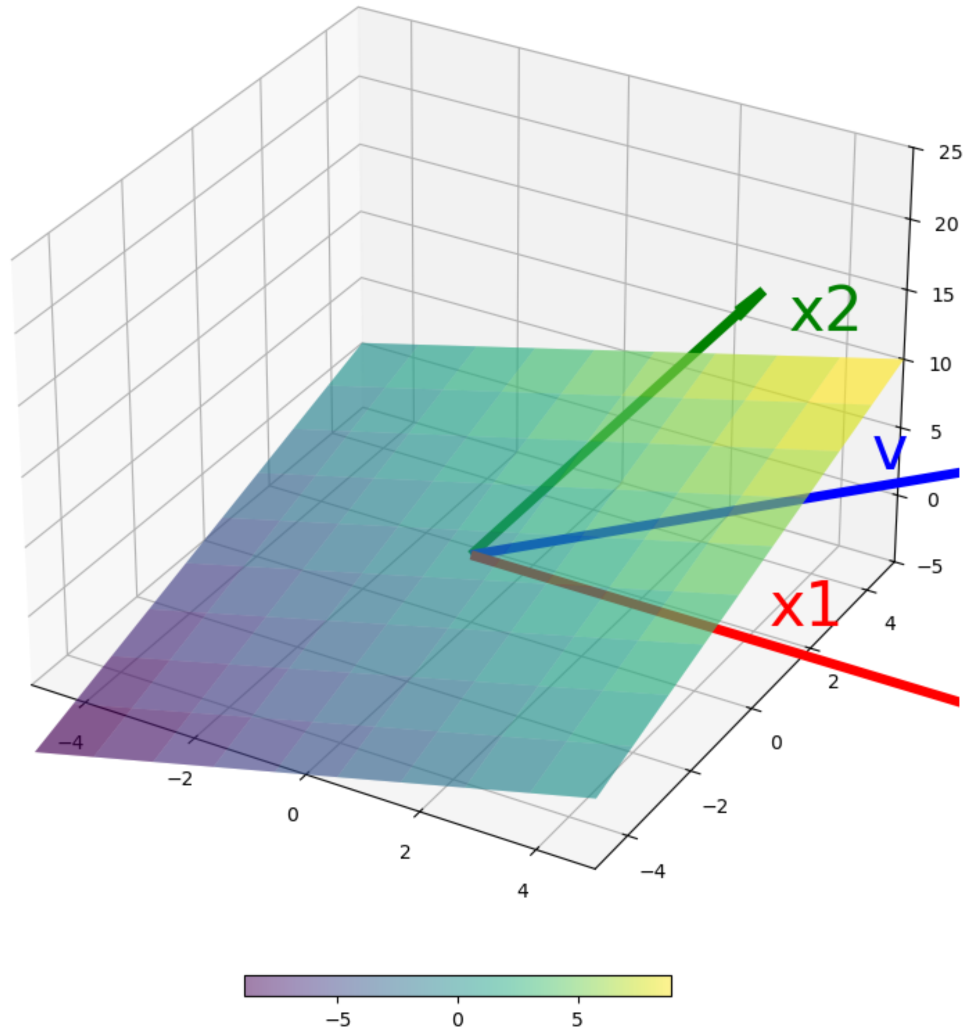
Practice

Function	Derivative
$y = x^n$	$y' = nx^{n-1}$

Common notation: x assumed to be vector

$$f(x) = 3x_1^2 + 4x_1x_2^2 + 2x_1 + 4x_3$$

Doesn't have a 'bottom'



$$\nabla f(x, y) = 0?$$

Impossible, no local minimum

$$\frac{\partial f}{\partial x_3} = \qquad 4$$

Now to vector valued functions

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}$$

$f_i(x)$ are scalar functions with gradients

Stack them up!

$$\nabla f(x) = \begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \end{bmatrix}$$

Vector-valued function

$$f(x_1, x_2) = \begin{bmatrix} 3x_1^2 + 5x_1x_2 \\ x_2^2 + 3x_1 \sin(x_2) \end{bmatrix}$$

Matrix, tensor-valued function…

Now to vector valued functions

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}$$

Example

$$f(x_1, x_2) = \begin{bmatrix} 3x_1^2 + 5x_1x_2 \\ x_2^2 + 3x_1\sin(x_2) \end{bmatrix}$$

$f_i(x)$ are scalar functions with gradients

Stack them up!

$$\nabla f(x) = \begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \end{bmatrix}$$

$$\nabla f(x) = \begin{bmatrix} \nabla f_1(x) = [6x_1 + 5x_2, \quad 5x_1] \\ \nabla f_2(x) = [3\cos(x_2), \quad 2x_2 + 3x_1\cos(x_2)] \end{bmatrix}$$

The Jacobian is the gradient for vector-valued functions

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix}$$

$$x = [x_1, x_2, \ldots x_m]$$

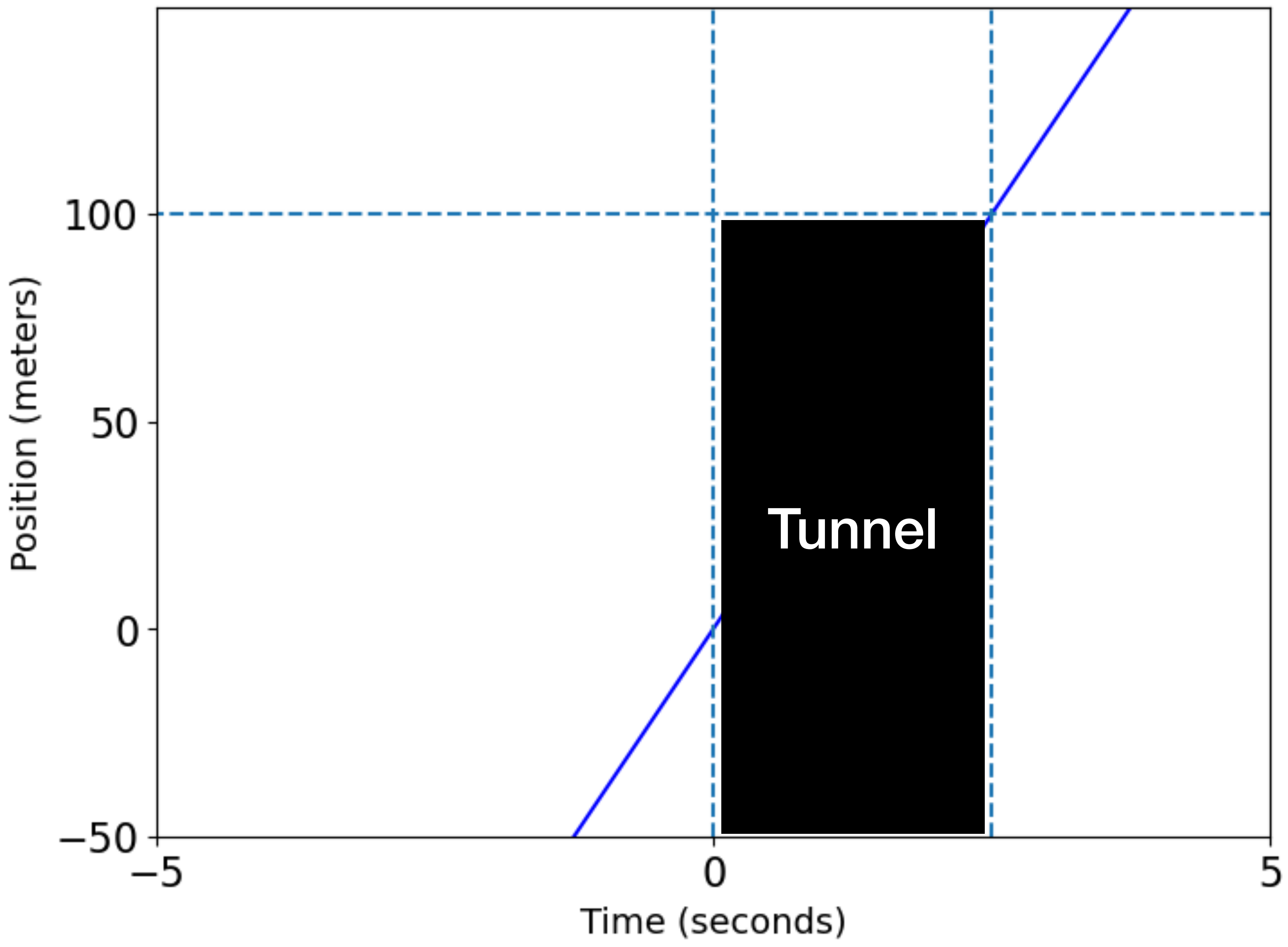
Stack them up!

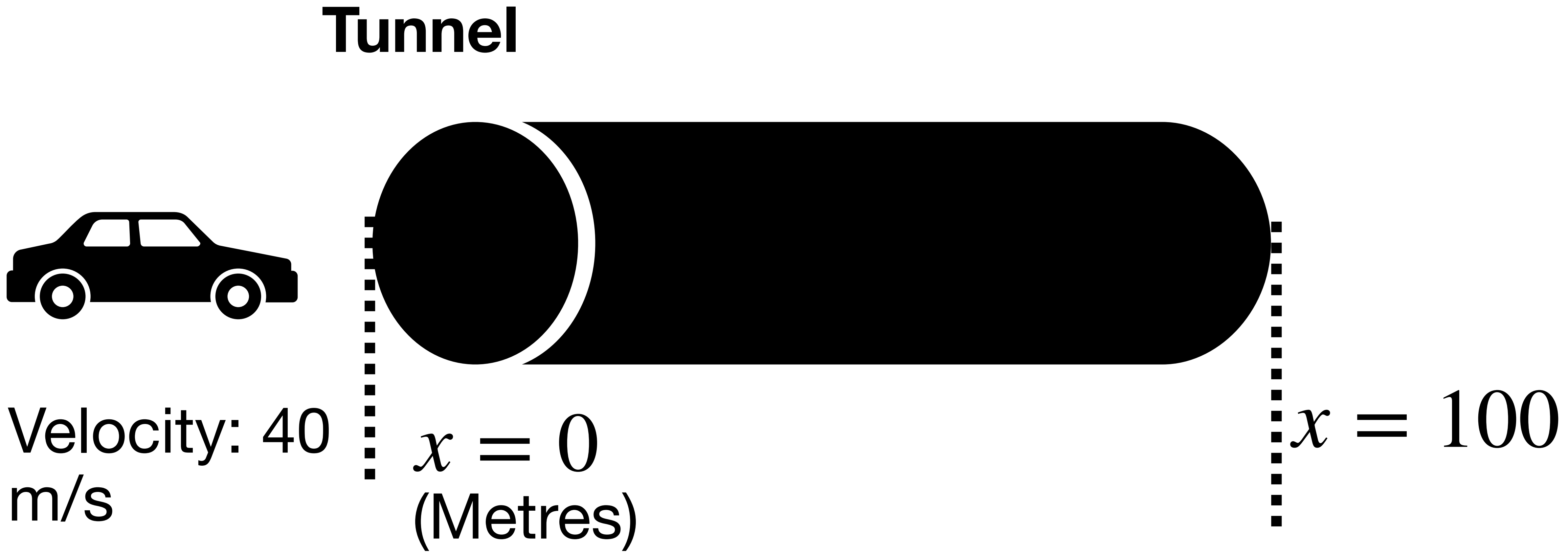
$$\nabla f(x) = \begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \end{bmatrix}$$

The Jacobian matrix

$$\begin{array}{cc} & m \textbf{ columns} \\ \nabla f(x) = & \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2}, \ldots, & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2}, \ldots, & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2}, \ldots, & \frac{\partial f_n}{\partial x_m} \end{bmatrix} & n \textbf{ rows} \end{array}$$

Taylor series: calculus for extrapolation





Measured just before entry

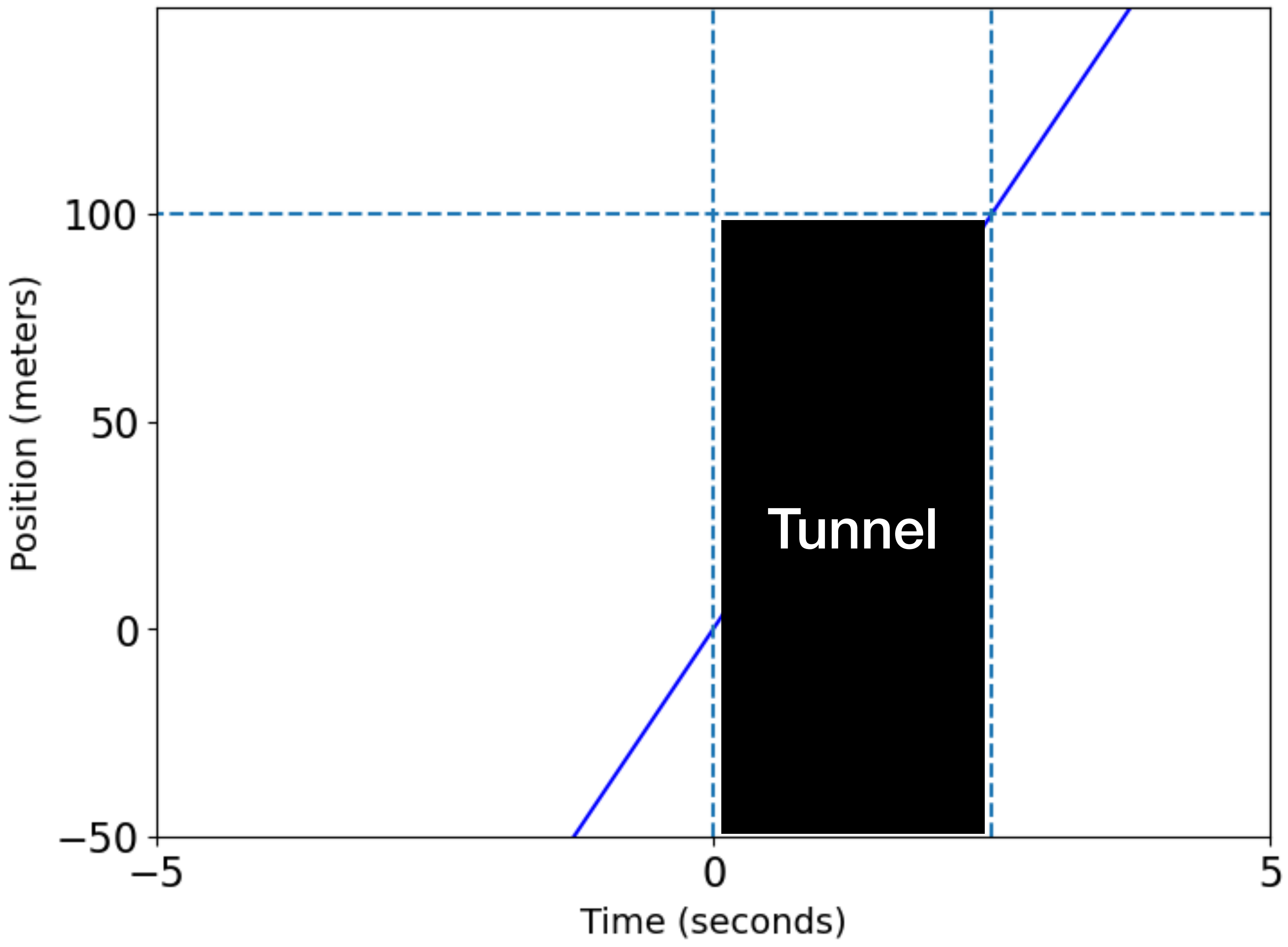
Estimate position after one second?

40 metres

Why might this be wrong?

Speed changes in tunnel

Taylor series: calculus for extrapolation



Linear approximation of position in tunnel

$$x(\Delta t) \approx b\Delta t + c$$

What are b and c?

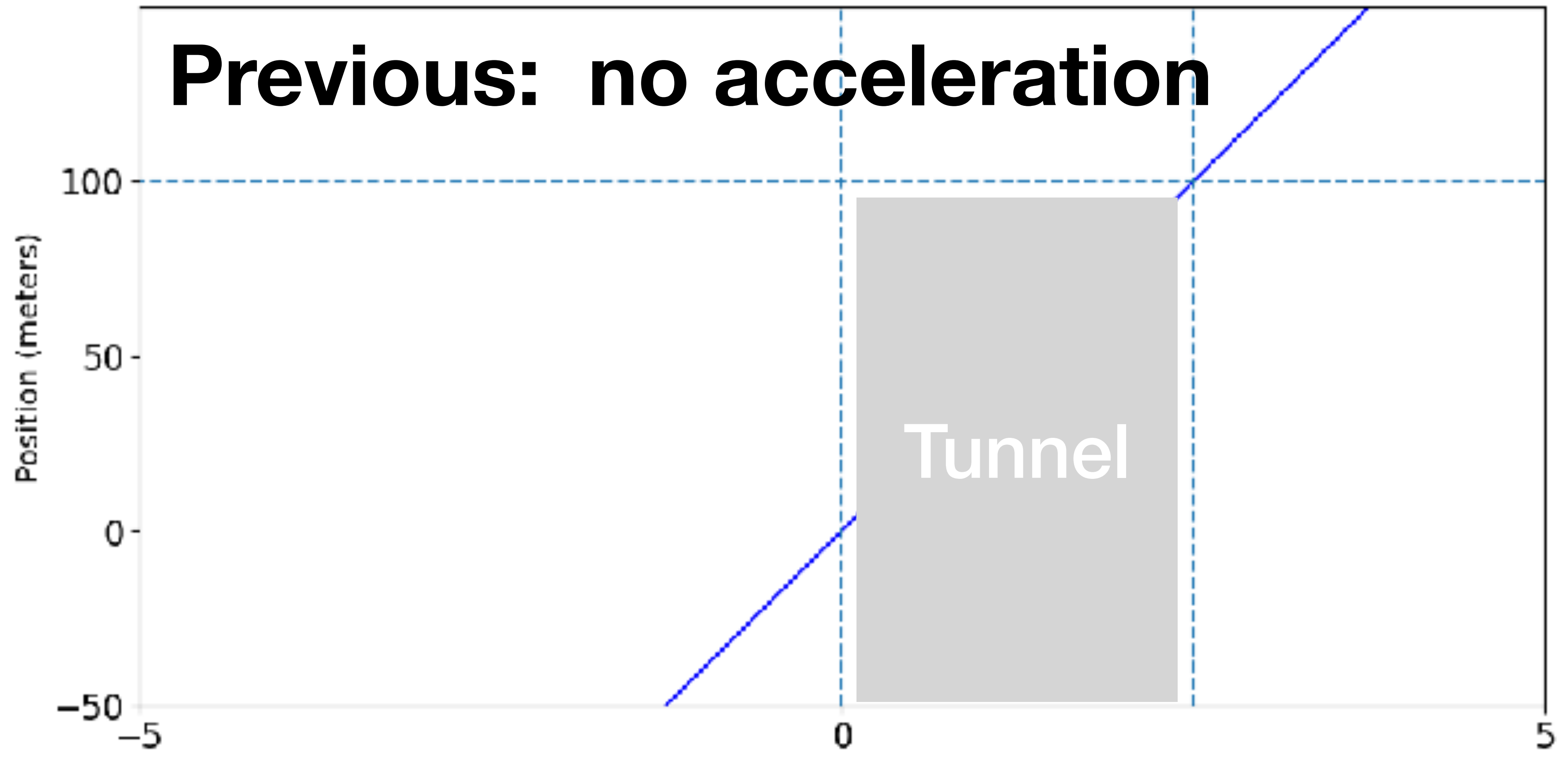
$c = x(0)$

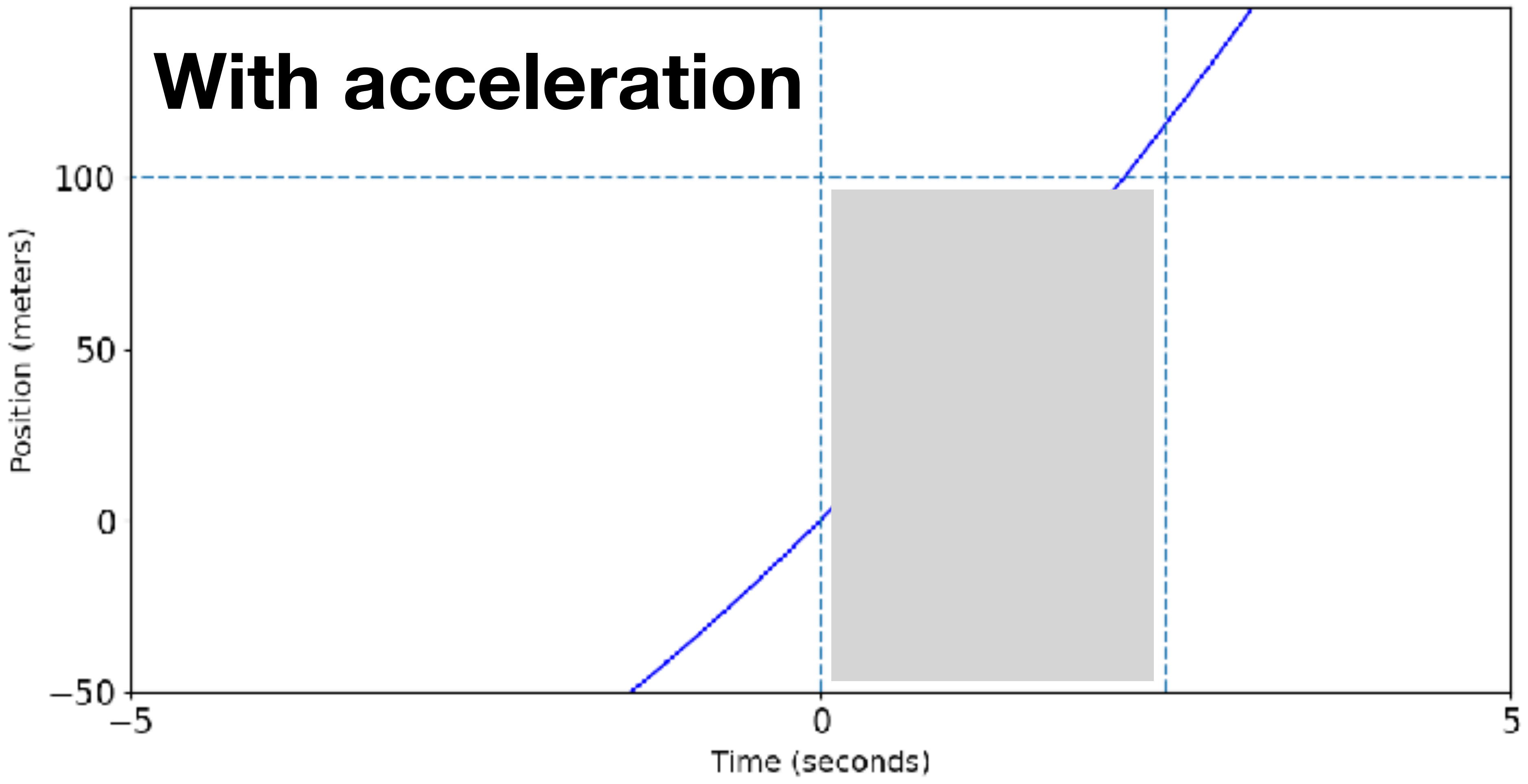
$b = x'(\Delta t)$ (differentiated both sides)

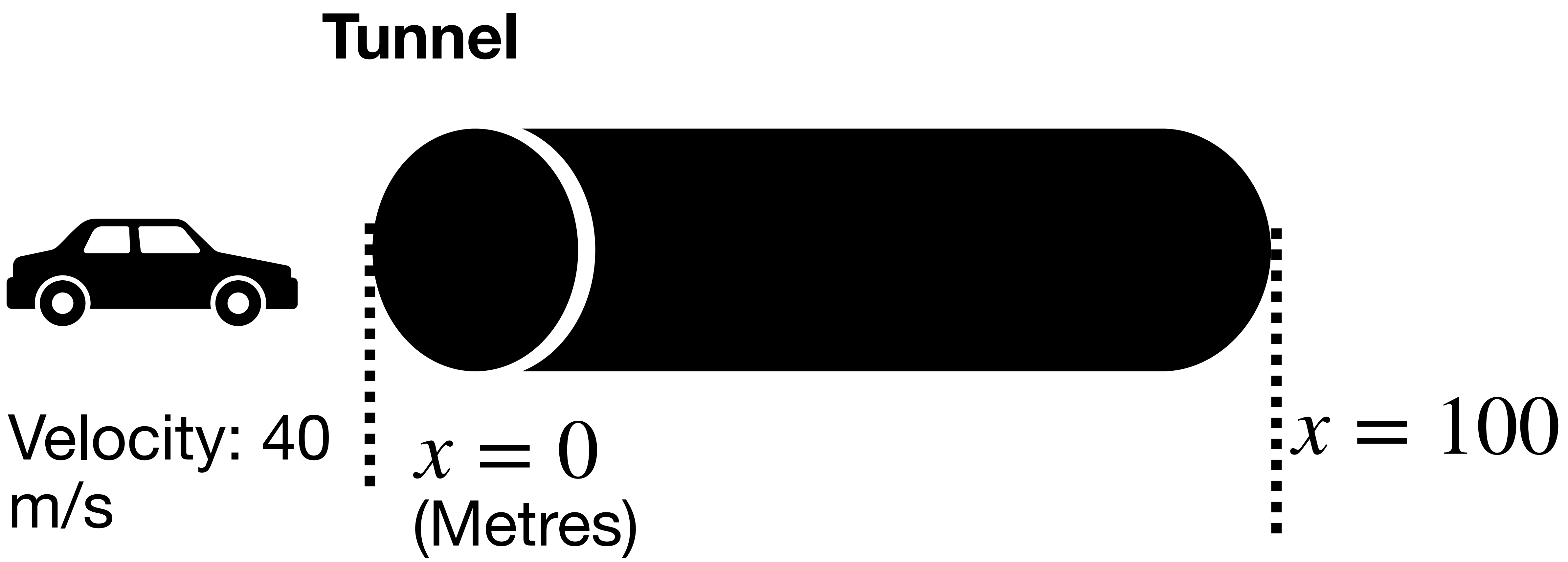
$\approx x'(0)$ Can't measure in tunnel!

$$x(\Delta t) \approx \underbrace{x(0)}_{c} + \underbrace{x'(0)\Delta t}_{b\Delta t}$$

Factor in acceleration?







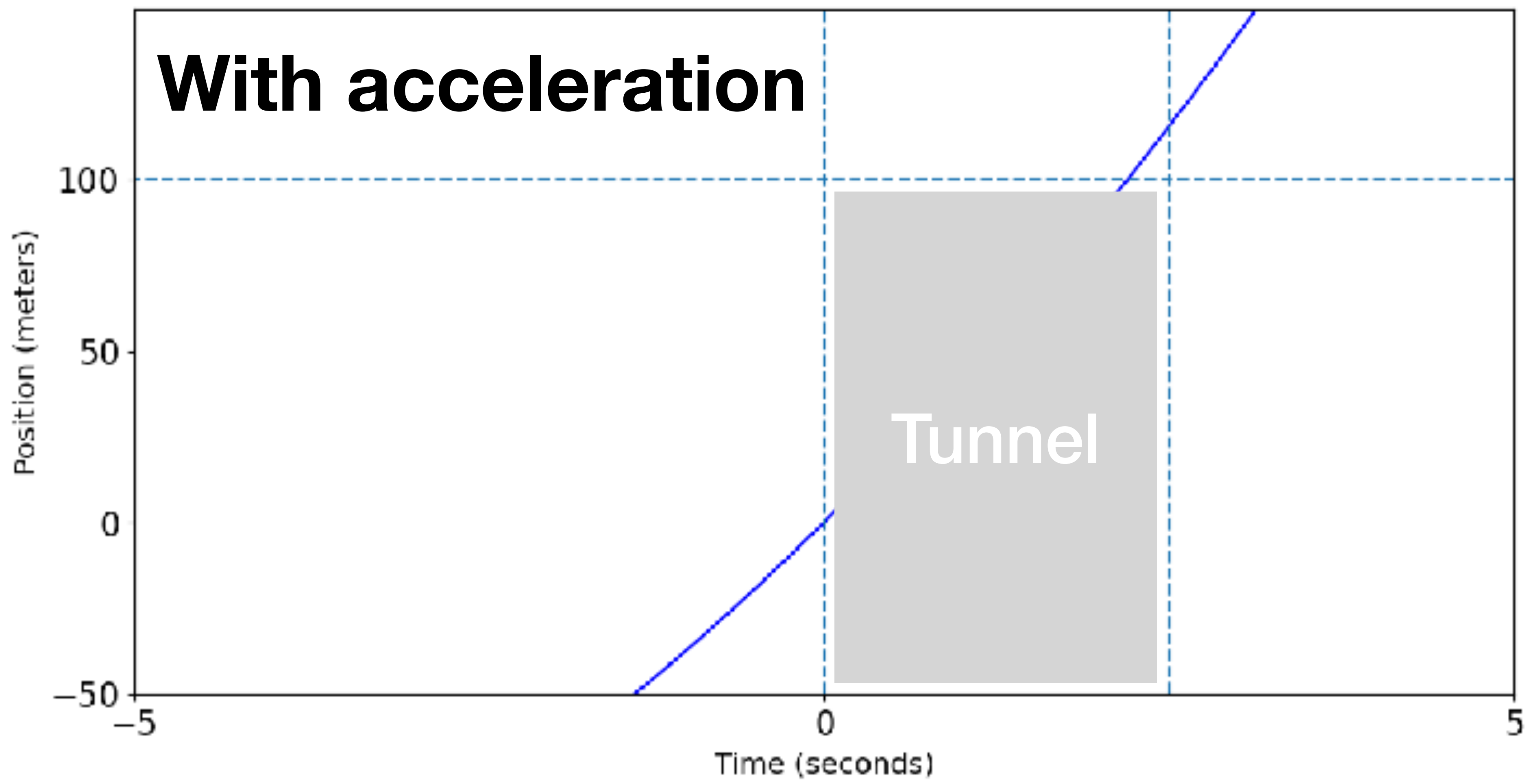
Measured just before entry

Estimate position after one second?

Clears tunnel faster!

Taylor series: calculus for extrapolation

:= means "we are defining as"



$$x(\Delta t) \approx x(0) + x'(0)\Delta t + \frac{1}{2}x''(0)(0)^2$$

c $b\Delta t$ $a\Delta t^2$

Quadratic approximation of position in tunnel

$$x(\Delta t) \approx a(\Delta t)^2 + b\Delta t + c$$

What are a,b and c?

From before

$c = x(0)$

$b \approx x'(0)$

(differentiated both sides)

$2a = x''(\Delta t)$

(differentiated both sides twice)

$\approx x''(0)$

Taylor series: calculus for extrapolation

1st order Taylor expansion

Linear approximation of position in tunnel (velocity)

$$x(\Delta t) \approx b\Delta t + c$$

2nd order Taylor expansion

Quadratic approximation of position in tunnel (+ accel)

$$x(\Delta t) \approx a(\Delta t)^2 + b\Delta t + c$$

3rd order Taylor expansion (rare)

$$x(\Delta t) \approx a(\Delta t)^3 + b(\Delta t)^2 + c\Delta t + d$$

Higher order Taylor expansion (rare)…

0-order Taylor expansion?

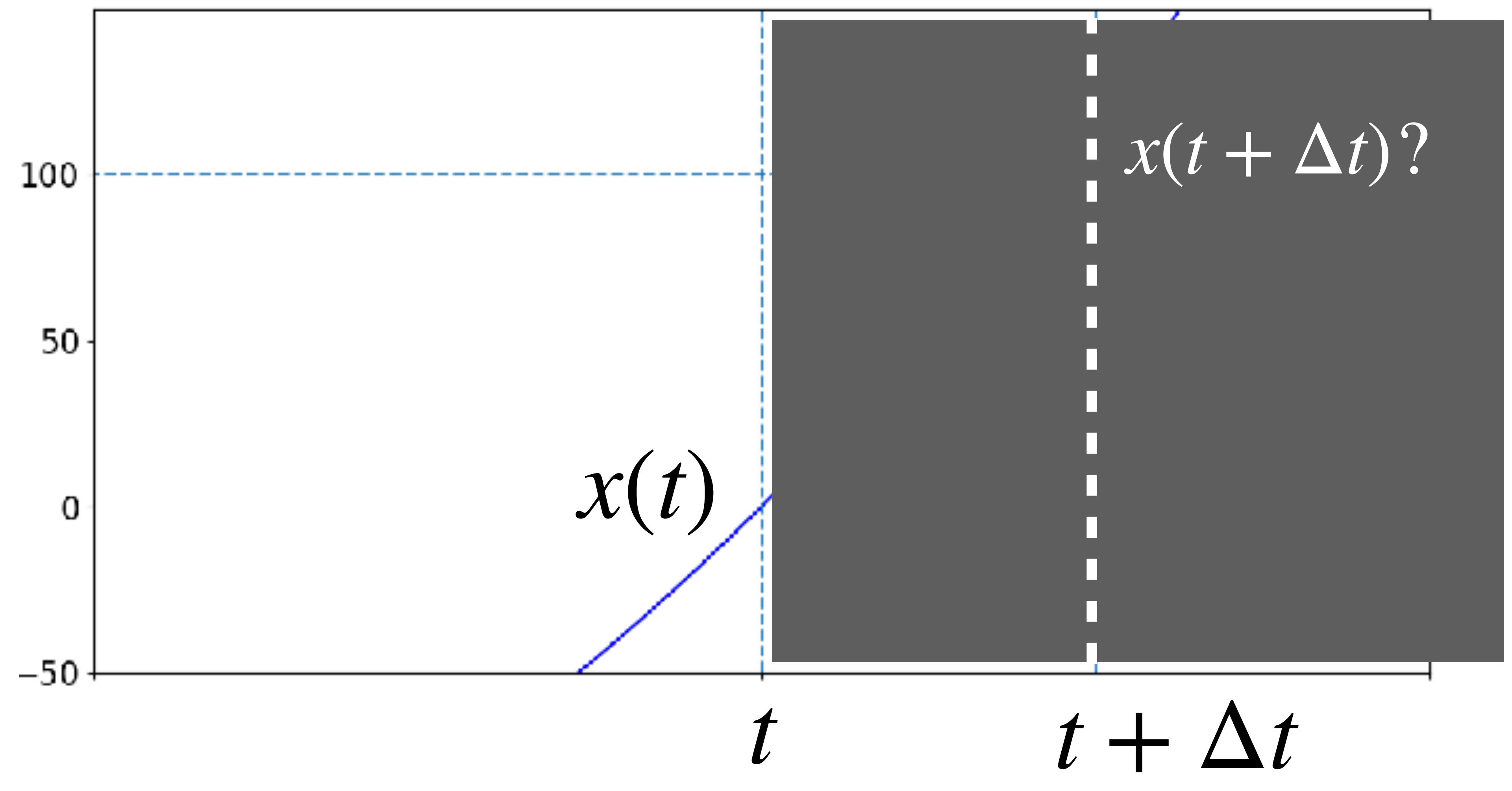
$$x(\Delta t) \approx c = x(0)$$

"Car is at the tunnel entrance"
(ignore velocity)

Taylor series: general formula

$$x(t+\Delta t) \approx x(t) \quad + x'(t)\Delta t \quad + \frac{1}{2}x''(t)(\Delta t)^2 \quad +\text{h.o.t}$$

0-order term | 1st order term | 2nd order term | "Higher order terms"



How would a function change if I changed the independent variable (t)?

Know: derivatives at current value $x'(t), x''(t), \ldots$

Why useful?

Extrapolating how a function changes is key to optimisation algorithms

Recall this next term!
(Backpropagation)

Will use later in course to understand optimisation algorithms

Neural network?
Objective function?

How would a ~~function~~ change if I changed the independent variable?

Weights?
Optimisation variables?

Know: derivatives at current value

$x'(t), x''(t), \ldots$