Math Basics for DS

Practical Session 9

Today

- Short quiz univariate calculus
- Multivariate calculus

Quiz

- https://forms.gle/hV6QkGjuC8RZgRMp6

Partial Derivatives

- Derivative for univariate functions:

$$y = f(x), \qquad \frac{dy}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

Partial Derivatives

- Derivative for univariate functions:

$$y = f(x), \qquad \frac{dy}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

- **Partial derivative** for multivariate functions:

$$y = f(x_1, \dots, x_n), \qquad \frac{\partial y}{\partial x_i} = f'_{x_i} = \lim_{\Delta x \to 0} \frac{f(x_1, \dots, x_i + \Delta x_i, \dots, x_n) - f(x_1, \dots, x_i, \dots x_n)}{\Delta x_i}$$

Compute derivative with respect to x_i
regarding all other variables as constants.

girafe ai

Partial Derivatives - Example

- Let $f(x, y) = x^3 + x^2y^3 - 2y^2$.

- Find $f_x(2,1)$ and $f_y(2,1)$.

Partial Derivatives - Example

- Let $f(x, y) = x^3 + x^2y^3 - 2y^2$.
- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x, y) = \frac{\partial f}{\partial x} =$$

$$f_y(x, y) = \frac{\partial f}{\partial y} =$$

Partial Derivatives - Example

- Let $f(x, y) = x^3 + x^2y^3 - 2y^2$.

- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x, y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3$$

$$f_y(x, y) = \frac{\partial f}{\partial y} =$$

girafe ai

Partial Derivatives - Example

- Let $f(x, y) = x^3 + x^2y^3 - 2y^2$.
- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x, y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3 \Big|_{(2,1)} =$$

$$f_y(x, y) = \frac{\partial f}{\partial y} =$$

Partial Derivatives - Example

- Let $f(x, y) = x^3 + x^2y^3 - 2y^2$.

- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x, y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3 \Big|_{(2,1)} = 12 + 4 = 16$$

$$f_y(x, y) = \frac{\partial f}{\partial y} =$$

Partial Derivatives - Example

- Let $f(x,y) = x^3 + x^2y^3 - 2y^2$.
- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x,y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3 \Big|_{(2,1)} = 12 + 4 = 16$$

$$f_y(x,y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y$$

girafe ai

Partial Derivatives - Example

- Let $f(x,y) = x^3 + x^2y^3 - 2y^2$.
- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x,y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3 \Big|_{(2,1)} = 12 + 4 = 16$$

$$f_y(x,y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y \Big|_{(2,1)} =$$

girafe ai

Partial Derivatives - Example

- Let $f(x, y) = x^3 + x^2y^3 - 2y^2$.
- Find $f_x(2,1)$ and $f_y(2,1)$.

$$f_x(x, y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3 \Big|_{(2,1)} = 12 + 4 = 16$$

$$f_y(x, y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y \Big|_{(2,1)} = 12 - 4 = 8$$

The Chain Rule

- For univariate functions:

$$\left(f\left(g(x)\right)\right)' = f'\left(g(x)\right) \cdot g'(x) \quad \Leftrightarrow \quad \frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$$

The Chain Rule

- For univariate functions:

$$\big(f(g(x))\big)' = f'(g(x)) \cdot g'(x) \quad \Leftrightarrow \quad \frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$$

- For multivariate functions:

$$u(x_1, \dots, x_n), \qquad x_i = x_i(t_1, \dots, t_m)$$

$$\frac{\partial u}{\partial t_k} = \frac{\partial u}{\partial x_1} \cdot \frac{\partial x_1}{\partial t_k} + \frac{\partial u}{\partial x_2} \cdot \frac{\partial x_2}{\partial t_k} + \cdots + \frac{\partial u}{\partial x_n} \cdot \frac{\partial x_n}{\partial t_k}, \qquad k = 1, \dots, m$$

The Chain Rule - Example

$$u = x^2 y + 3xy^4, \qquad x = \sin 2t\,, \qquad y = \cos t$$

The Chain Rule - Example

$$u = x^2y + 3xy^4, \qquad x = \sin 2t\,, \qquad y = \cos t$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial t}$$

The Chain Rule - Example

$$u = x^2 y + 3xy^4, \qquad x = \sin 2t\,, \qquad y = \cos t$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial t}$$

$$\frac{\partial u}{\partial t} =$$

The Chain Rule - Example

$$u = x^2 y + 3xy^4, \qquad x = \sin 2t\,, \qquad y = \cos t$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial t}$$

$$\frac{\partial u}{\partial t} =$$

The Chain Rule - Example

$$u = x^2 y + 3xy^4, \qquad x = \sin 2t\,, \qquad y = \cos t$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial t}$$

$$\frac{\partial u}{\partial t} = (2xy + 3y^4) \cdot 2\cos 2t +$$

The Chain Rule - Example

$$u = x^2 y + 3xy^4, \qquad x = \sin 2t, \qquad y = \cos t$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial t}$$

$$\frac{\partial u}{\partial t} = (2xy + 3y^4) \cdot 2\cos 2t + (x^2 + 12xy^3)(-\sin t)$$

girafe ai

The Chain Rule - Example

$$u = x^4 y + y^2 z^3$$

$$x = rse^t, \qquad y = rs^2 e^{-t}, \qquad z = r^2 s \sin t$$

The Chain Rule - Example

$$u = x^4 y + y^2 z^3$$

$$x = rse^t, \qquad y = rs^2 e^{-t}, \qquad z = r^2 s \sin t$$

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial s} + \frac{\partial u}{\partial z} \cdot \frac{\partial z}{\partial s}$$

The Chain Rule - Example

$$u = x^4 y + y^2 z^3$$

$$x = rse^t, \qquad y = rs^2 e^{-t}, \qquad z = r^2 s \sin t$$

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial s} + \frac{\partial u}{\partial z} \cdot \frac{\partial z}{\partial s}$$

$$\frac{\partial u}{\partial s} =$$

girafe ai

The Chain Rule - Example

$$u = x^4 y + y^2 z^3$$

$$x = rse^t, \qquad y = rs^2 e^{-t}, \qquad z = r^2 s \sin t$$

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial s} + \frac{\partial u}{\partial z} \cdot \frac{\partial z}{\partial s}$$

$$\frac{\partial u}{\partial s} = 4x^3 y \cdot re^t + (x^4 + 2yz^3) \cdot 2rse^{-t} + 3y^2 z^2 \cdot r^2 \sin t$$

Gradient

- A vector of partial derivatives:

$$\nabla f(x_1, \dots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

Gradient

- A vector of partial derivatives:

$$\nabla f(x_1, \dots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)$$

- Example:

$$f(x, y, z) = e^{xy} \ln z$$

Gradient

- A vector of partial derivatives:

$$\nabla f(x_1, \ldots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

- Example:

$$f(x, y, z) = e^{xy} \ln z$$

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) =$$

Gradient

- A vector of partial derivatives:

$$\nabla f(x_1, \dots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

- Example:

$$f(x, y, z) = e^{xy} \ln z$$

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = \left(y e^{xy} \ln z, \quad x e^{xy} \ln z, \quad \frac{e^{xy}}{z} \right).$$

Gradient

Gradient shows the direction of the maximal growth of the function.

Negative gradient = direction of the maximal descent of the function.

Higher Derivatives

- Consider $f(x, y)$.
- $f'_x(x, y),\ f'_y(x, y)$ – also functions of two variables.
  We can compute their partial derivatives!

girafe ai

Higher Derivatives

- Consider $f(x, y)$.

- $f'_x(x, y),\ f'_y(x, y)$ – also functions of two variables.
  We can compute their partial derivatives!

$$(f'_x)'_x = f''_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2}, \qquad (f'_x)'_y = f''_{xy}(x, y) = \frac{\partial^2 f}{\partial y \partial x}$$

$$\left(f'_y\right)'_y = f''_{yy}(x, y) = \frac{\partial^2 f}{\partial y^2}, \qquad \left(f'_y\right)'_x = f''_{yx}(x, y) = \frac{\partial^2 f}{\partial x \partial y}$$

girafe ai

Higher Derivatives - Example

- Consider $f(x,y) = x^3 + x^2y^3 - 2y^2$.

$$f_x(x,y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3, \qquad f_y(x,y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y$$

$$f''_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = \qquad , \qquad f''_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} =$$

$$f''_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = \qquad , \qquad f''_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y} =$$

girafe ai

Higher Derivatives - Example

- Consider $f(x,y) = x^3 + x^2y^3 - 2y^2$.

$$f_x(x,y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3, \qquad f_y(x,y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y$$

$$f''_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = 6x + 2y^3, \qquad f''_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} =$$

$$f''_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = \qquad\qquad , \qquad f''_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y} =$$

girafe ai

Higher Derivatives - Example

- Consider $f(x,y) = x^3 + x^2y^3 - 2y^2$.

$$f_x(x,y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3, \qquad f_y(x,y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y$$

$$f''_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = 6x + 2y^3, \qquad f''_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} = 6x$$

$$f''_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = \qquad\qquad , \qquad f''_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y} =$$

girafe ai

Higher Derivatives - Example

- Consider $f(x,y) = x^3 + x^2y^3 - 2y^2$.

$$f_x(x,y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3, \qquad f_y(x,y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y$$

$$f''_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = 6x + 2y^3, \qquad f''_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} = 6x$$

$$f''_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = 6x^2y - 4, \qquad f''_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y} =$$

Higher Derivatives - Example

- Consider $f(x, y) = x^3 + x^2y^3 - 2y^2$.

$$f_x(x, y) = \frac{\partial f}{\partial x} = 3x^2 + 2xy^3, \qquad f_y(x, y) = \frac{\partial f}{\partial y} = 3x^2y^2 - 4y$$

$$f''_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2} = 6x + 2y^3, \qquad f''_{xy}(x, y) = \frac{\partial^2 f}{\partial y \partial x} = 6x$$

$$f''_{yy}(x, y) = \frac{\partial^2 f}{\partial y^2} = 6x^2y - 4, \qquad f''_{yx}(x, y) = \frac{\partial^2 f}{\partial x \partial y} = 6x$$

Hessian

- A matrix of second derivatives.
- $f(x_1, \dots, x_n)$

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 x_1} & \cdots & \frac{\partial^2 f}{\partial x_n x_1} \\ \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Hessian - Example

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

$$f''_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = 6x + 2y^3, \qquad f''_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} = 6x$$

$$f''_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = 6x^2y - 4, \qquad f''_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y} = 6x$$

$$H = \begin{pmatrix} \quad \end{pmatrix}$$

Hessian - Example

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

$$f''_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = 6x + 2y^3, \qquad f''_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} = 6x$$

$$f''_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = 6x^2y - 4, \qquad f''_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y} = 6x$$

$$H = \begin{pmatrix} 6x + 2y^3 & 6x \\ 6x & 6x^2y - 4 \end{pmatrix}$$

Extrema

- Univariate case: a stationary point x_0 is a local minimum (maximum) if

$$f''(x_0) > 0 \ (f''(x_0) < 0).$$

- Multivariate case:

x_0 is a stationary point: $\nabla f(x_0) = 0$

$H -$ Hessian.

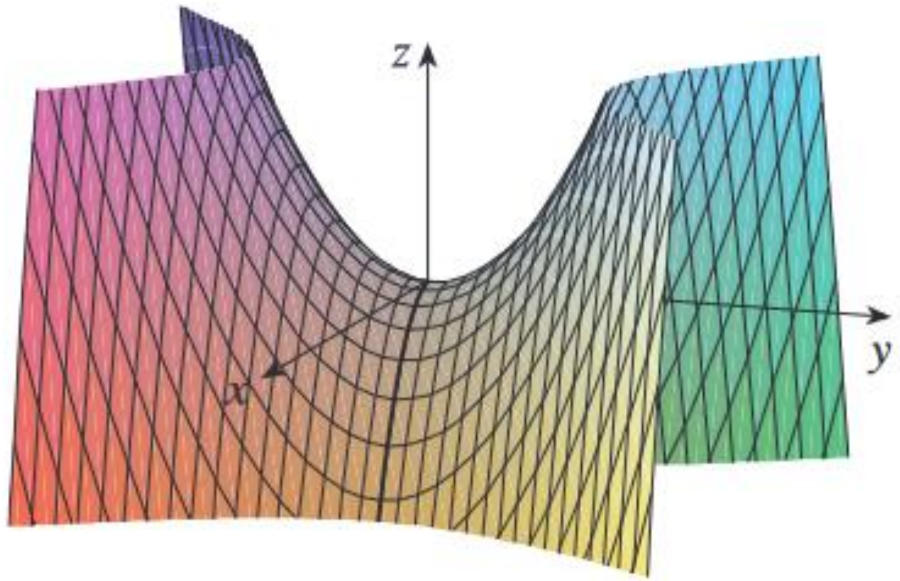
Then, if H is a **positive-definite** matrix, then x_0 is a **local minimum**.

If H is a **negative-definite** matrix, then x_0 is a **local maximum**.

If $\det H = 0$, we need to check manually.

Otherwise, x_0 is a **saddle point**.

Saddle Points



Positive vs Negative Definite Matrices

- A matrix A is called positive-definite if

$$x^T A x > 0 \ \forall x$$

- A matrix A is called negative-definite if

$$x^T A x < 0 \ \forall x$$

Positive vs Negative Definite Matrices

- How to check if a matrix is positive (negative) definite?

Positive vs Negative Definite Matrices

- How to check if a matrix is positive (negative) definite?

  Check principal minors D_k!

For positive definite:

$$D_1 > 0, \quad D_2 > 0, \quad \ldots, D_n > 0$$

For negative definite:

$$D_1 < 0, \quad D_2 > 0, \quad D_3 < 0, \ldots$$

$$H = \begin{pmatrix} \boxed{\dfrac{\partial^2 f}{\partial x_1^2}} & \dfrac{\partial^2 f}{\partial x_2 x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n x_1} \\ \dfrac{\partial^2 f}{\partial x_1 x_2} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n x_1} & \dfrac{\partial^2 f}{\partial x_n x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) =$$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $\nabla f = 0 \quad \Leftrightarrow$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f_x', f_y'\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $\nabla f = 0 \quad \Leftrightarrow$

$$\begin{aligned} 4x^3 - 4y &= 0 \\ 4y^3 - 4x &= 0 \end{aligned} \quad \Leftrightarrow$$

girafe ai

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $\nabla f = 0 \quad \Longleftrightarrow$

$$\begin{matrix} 4x^3 - 4y = 0 \\ 4y^3 - 4x = 0 \end{matrix} \quad \Longleftrightarrow \quad \begin{matrix} y = x^3 \\ x^9 - x = 0 \end{matrix} \quad \Longleftrightarrow$$

girafe ai

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f_x', f_y'\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $\nabla f = 0 \quad \Leftrightarrow$

$$\begin{matrix} 4x^3 - 4y = 0 \\ 4y^3 - 4x = 0 \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} y = x^3 \\ x^9 - x = 0 \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} y = x^3 \\ x(x^8 - 1) = 0 \end{matrix}$$

Extrema – Example

$$f(x,y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $\nabla f = 0 \quad \Leftrightarrow$

$$\begin{matrix} 4x^3 - 4y = 0 \\ 4y^3 - 4x = 0 \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} y = x^3 \\ x^9 - x = 0 \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} y = x^3 \\ x(x^8 - 1) = 0 \end{matrix}$$

Stationary points: $(0,0),\ (-1,-1),\ (1,1)$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0,0)$, $(-1,-1)$, $(1,1)$

$$H = \left(\qquad\qquad \right)$$

Extrema – Example

$$f(x,y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f_x', f_y'\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0,0),\ (-1,-1),\ (1,1)$

$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

girafe ai

Extrema – Example

$$f(x,y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0,0),\ (-1,-1),\ (1,1)$

$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

$\det H = 144x^2y^2 - 16|_{(0;0)} < 0,\ (0,0)$ – saddle point.

girafe ai

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0, 0),\ (-1, -1),\ (1, 1)$

$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

$$\det H = 144x^2y^2 - 16 \qquad , \qquad \det H_1 = 12x^2$$

girafe ai

Extrema – Example

$$f(x,y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0,0),\ (-1,-1),\ (1,1)$

$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

$$\det H = 144x^2y^2 - 16 \Big|_{(1;1)} > 0, \qquad \det H_1 = 12x^2 \Big|_{(1,1)} > 0$$

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0, 0),\ (-1, -1),\ (1, 1)$

$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

$$\det H = 144x^2y^2 - 16\Big|_{(1;1)} > 0, \qquad \det H_1 = 12x^2\Big|_{(1,1)} > 0$$

$(1, 1)$ – local minimum.

girafe ai

Extrema – Example

$$f(x,y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f'_x, f'_y\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0,0),\ (-1,-1),\ (1,1)$

$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

$$\det H = 144x^2y^2 - 16 \qquad\qquad , \qquad \det H_1 = 12x^2$$

girafe ai

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$

$$\nabla f = \left(f_x', f_y'\right) = (4x^3 - 4y, \qquad 4y^3 - 4x)$$

Stationary points: $(0, 0)$, $(-1, -1)$, $(1, 1)$

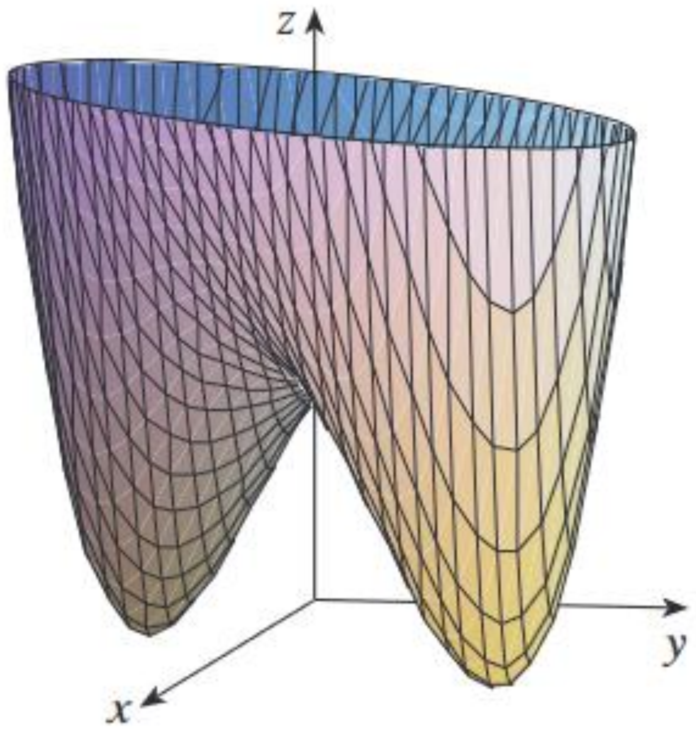
$$H = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}$$

$$\det H = 144x^2y^2 - 16 \Big|_{(-1;-1)} > 0, \qquad \det H_1 = 12x^2 \Big|_{(-1,-1)} > 0$$

$(-1, -1)$ – local minimum.

Extrema – Example

$$f(x, y) = x^4 + y^4 - 4xy + 1$$



Linear Regression the Other Way

Least Squares (Again)

- Remember the Least Squares problem:

$$\hat{y}_i = w_0 + w_1 x_1^i + \cdots + w_n x_n^i, \qquad y_i = \hat{y}_i + e_i, \qquad w = ?$$

girafe ai

Least Squares (Again)

- Remember the Least Squares problem:

$$\hat{y}_i = w_0 + w_1 x_1^i + \cdots + w_n x_n^i, \qquad y_i = \hat{y}_i + e_i, \qquad w = ?$$

- Before: solve by projecting y onto the $col(A)$:

$$\hat{w} = (X^T X)^{-1} X^T y$$

girafe ai

Least Squares (Again)

- Remember the Least Squares problem:

$$\hat{y}_i = w_0 + w_1 x_1^i + \cdots + w_n x_n^i, \qquad y_i = \hat{y}_i + e_i, \qquad w = ?$$

- Before: solve by projecting y onto the $col(A)$:

$$\hat{w} = (X^T X)^{-1} X^T y$$

- We can also see this as an optimization problem:

$$\mathcal{L}(w_0, \dots, w_n) = \sum e_i^2 = \sum (y_i - \hat{y})^2 = \sum \left(y_i - \left(w_0 + w_1 x_1^i + \cdots + w_n x_n^i\right)\right)^2 \to \min_{w_0, \dots, w_n}$$

girafe ai

Least Squares

- Let's consider a simple case: fitting a line to a set of points

$$(x_1, y_1), \qquad (x_2, y_2), \qquad \dots, \qquad (x_n, y_n)$$

$$y \approx w_0 + w_1 x$$

We want to pick the weights $w_0,\ w_1$
so that the sum of squares of the errors is the smallest possible:

$$\mathcal{L}(w_0, w_1) = \qquad \qquad \rightarrow \min_{w_0, w_1}$$

Least Squares

- Let's consider a simple case: fitting a line to a set of points

$$(x_1, y_1), \qquad (x_2, y_2), \qquad \dots , \qquad (x_n, y_n)$$

$$y \approx w_0 + w_1 x$$

We want to pick the weights $w_0,\ w_1$
so that the sum of squares of the errors is the smallest possible:

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

girafe ai

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \rightarrow \min_{w_0, w_1}$$

$$\frac{\partial \mathcal{L}}{\partial w_1} =$$

$$\frac{\partial \mathcal{L}}{\partial w_0} =$$

girafe ai

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

$$\frac{\partial \mathcal{L}}{\partial w_1} = -2 \sum_{i=1}^{n} x_i \cdot (y_i - w_0 - w_1 x_i) = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_0} = -2 \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i) = 0$$

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

$$\frac{\partial \mathcal{L}}{\partial w_1} = \sum_{i=1}^{n} x_i y_i - w_0 \sum_{i=1}^{n} x_i - w_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_0} = \sum_{i=1}^{n} y_i - w_1 \sum_{i=1}^{n} x_i - n w_0 = 0$$

girafe ai

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

$$w_0 \sum_{i=1}^{n} x_i + w_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

$$n w_0 + w_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

girafe ai

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

$$w_0 \sum_{i=1}^{n} x_i + w_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

$$w_0 + w_1 \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{n} \sum_{i=1}^{n} y_i \quad \Leftrightarrow \quad w_0 + w_1 \bar{x} = \bar{y}$$

girafe ai

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \rightarrow \min_{w_0, w_1}$$

$$w_0 \sum_{i=1}^{n} x_i + w_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

$$n w_0 + w_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$w_0 + w_1 \bar{x} = \bar{y}$$

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \rightarrow \min_{w_0, w_1}$$

$$n w_0 \sum_{i=1}^{n} x_i + n w_1 \sum_{i=1}^{n} x_i^2 = n \sum_{i=1}^{n} x_i y_i$$

$$w_0 + w_1 \bar{x} = \bar{y}$$

$$n w_0 \sum_{i=1}^{n} x_i + w_1 \left(\sum_{i=1}^{n} x_i \right)^2 = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

girafe ai

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

$$w_1 \left(\sum_{i=1}^{n} x_i \right)^2 - n w_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n \sum_{i=1}^{n} x_i y_i$$

$$w_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

Least Squares

$$\mathcal{L}(w_0, w_1) = \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \to \min_{w_0, w_1}$$

$$\widehat{w}_0 = \widehat{w}_1 \bar{x} = \bar{y},$$

$$\widehat{w}_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$