The Hessian is like an extension of the Jacobian

- Jacobian is made from first order derivatives into a vector
- Hessian group of second order derivatives into matrix.

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

Example: finding Hessian

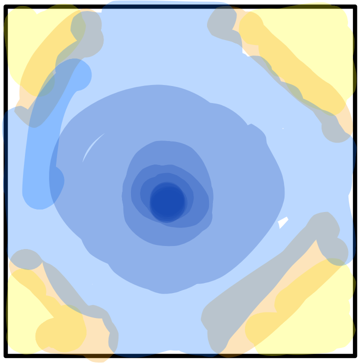
$$f(x, y, z) = x^2 y z$$

$$J = [2xyz, x^2 z, x^2 y]$$

$$H = \begin{bmatrix} 2yz & 2xz & 2xy \\ 2xz & 0 & x^2 \\ 2xy & x^2 & 0 \end{bmatrix}$$

- notice hessian matrix is symmetrical accross leading diagonal.
  ↳ if the function is continuous and not a step function.

Hessian: Case #1 & #2



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

$J = [2x, 2y]$

$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$|H| = 4$

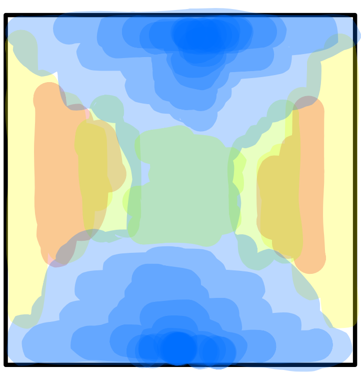
$J(0,0) = 0$, a gradient of zero.

If the hessian determinant is positive, the gradient of zero will either be a maximum or a minimum.

If the top left term of a hessian is positive then the gradient is a minimum. ? How?

↳ If it's negative $Ju = 0$ will be a maximum.

Hessian: Case #3 – A saddle point



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

$J = [2x, -2y]$

$H = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$

$|H| = -4$

$J(0,0) = 0$, a gradient of zero. is "flat".

Because hessian $|H|$ is negative this doesn't now max/min. Cause confusion in optimization.