Multi-Variable Calculus Notes

Table of Contents:

The Gradient:

- Suppose $f: R^{m \times n} \to R$ is a function that takes as input a matrix A of size $m \times n$ and returns a real value. Then, the gradient of f with respect to $A \in R^{m \times n}$ is the matrix of partial derivatives.

I.e. $$\nabla f(A) = \begin{bmatrix} \frac{\partial f(A)}{\partial A_{11}} & \frac{\partial f(A)}{\partial A_{12}} & \cdots & \frac{\partial f(A)}{\partial A_{1N}} \\ \vdots & & & \vdots \\ \frac{\partial f(A)}{\partial A_{m1}} & \cdots & \cdots & \frac{\partial f(A)}{\partial A_{mn}} \end{bmatrix}$$

- Find the gradient in each of the next examples.

E.g. 1 $f(x, y, z) = xy + z$

Soln:

$$\frac{\partial f}{\partial x} = y \quad \frac{\partial f}{\partial y} = x \quad \frac{\partial f}{\partial z} = 1$$

$$\therefore \nabla f(x, y, z) = [y, x, 1]$$

E.g. 2 $f(x, y, z) = xy + yz + xz$

Soln:

$$\frac{\partial f}{\partial x} = y + z, \quad \frac{\partial f}{\partial y} = x + z \quad \frac{\partial f}{\partial z} = y + x$$

$$\therefore \nabla f(x, y, z) = [y+z, x+z, x+y]$$

- Some properties of gradients:
  1. $\nabla(f(x) + g(x)) = \nabla f(x) + \nabla g(x)$
  2. For $t \in R$, $\nabla(t f(x)) = t(\nabla f(x))$

The Hessian:

- Suppose that $f: R^n \to R$ is a function that takes a vector in R^n and returns a real number. The Hessian matrix with respect to x, denoted as $\nabla_x^2 f(x)$ or H, is the nxn matrix of partial derivatives.

I.e.

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

- Find the Hessian matrix for each of the examples below:

E.g. 1 $f(x,y) = x^3 - 2xy - y^6$

Soln:

$f_x = 3x^2 - 2y$
$f_y = -2x - 6y^5$

$f_{xx} = 6x$
$f_{yy} = -30y^4$
$f_{xy} = f_{yx} = -2$

$$\therefore H = \begin{bmatrix} 6x & -2 \\ -2 & -30y^4 \end{bmatrix}$$

E.g. 2 $f(x,y) = y^4 + x^3 + 3x^2 + 4y^2 - 4xy - 5y + 8$

Soln:

$f_x = 3x^2 + 6x - 4y$

$f_y = 4y^3 + 8y - 4x - 5$

$f_{xx} = 6x + 6$

$f_{yy} = 12y^2 + 8$

$f_{xy} = f_{yx} = -4$

$$\therefore H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 6x+6 & -4 \\ -4 & 12y^2+8 \end{bmatrix}$$

E.g. 3 $f(x,y) = x^2y + y^2x$

Soln:

$$\left.\begin{aligned} f_x &= 2xy + y^2 \\ f_y &= x^2 + 2yx \\ f_{xx} &= 2y \\ f_{yy} &= 2x \\ f_{xy} &= f_{yx} = 2x + 2y \end{aligned}\right\} H = \begin{bmatrix} 2y & 2x+2y \\ 2x+2y & 2x \end{bmatrix}$$

- Note: The Hessian matrix is always symmetrical.

Gradients and Hessians of Quadratic and Linear Functions:
- $\nabla b^T x = b$
- $\nabla x^T A x = 2Ax$ if A is symmetric
- $\nabla^2 x^T A x = 2A$ if A is symmetric

Least Squares
- Let $A \in R^{m \times n}$ have a full rank.
  Let $b \in R^m$ s.t. $b \notin R(A)$
  In this situtation, we want to find a vector $x \in R^n$ s.t. Ax is as close to b as possible, as measured by the square of the Euclidean norm $(\|Ax-b\|_2)^2$.

  Using the fact $(\|x\|_2)^2 = x^T x$, we have:

$$(\|Ax-b\|_2)^2 = (Ax-b)^T (Ax-b)$$
$$= x^T A^T x - 2b^T Ax + b^T b$$

Taking the gradient w.r.t x, we have:

$$\nabla_x (x^T A^T x - 2b^T Ax + b^T b)$$
$$= \nabla_x (x^T A^T x) - \nabla_x (2b^T Ax) + \nabla_x (b^T b)$$
$$= 2A^T Ax - 2A^T b$$

Setting the last expression to 0, and solving for x, we get:

$$x = (A^T A)^{-1} A^T b$$

Gradients of the Determinant:

- Let $A \in \mathbb{R}^{n \times n}$. We want to find $\nabla_A |A|$.

$$\frac{\partial}{\partial A_{k\ell}} |A| = \frac{\partial}{\partial A_{k\ell}} \sum_{i=1}^{n} (-1)^{i+j} A_{ij} |A_{\setminus i \setminus j}|$$

$$= (-1)^{k+\ell} |A_{\setminus k \setminus \ell}|$$

$$= (\mathrm{adj}(A))_{\ell k}$$

$$= (\mathrm{adj}(A))^T$$

$$= |A| A^{-T}$$

Eigenvalues as Optimization:

- If we want to optimize (min or max) $f(x,y,z)$ subject to the constraint $g(x,y,z) = k$, we can do:
  1. $\nabla f(x,y,z) = \lambda \nabla g(x,y,z)$
  2. $g(x,y,z) = k$
  3. Plug all solns into $f(x,y,z)$ and find the max and min.

- $L(x, \lambda) = f(x) - \lambda g(x)$ is called Lagrange function.
- λ is called Lagrange Multiplier.
- Note: $\nabla g \neq 0$ at the point

- E.g.

1. Find the max and min of $f(x,y) = 5x - 3y$ subject to the constraint $x^2 + y^2 = 136$.

Soln:

$\nabla f(x,y) = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$ $\lambda \nabla g(x,y) = \begin{bmatrix} 2\lambda x \\ 2\lambda y \end{bmatrix}$

$5 = 2\lambda x$

$-3 = 2\lambda y$

$x^2 + y^2 = 136$

$x = \dfrac{5}{2\lambda}$, $y = \dfrac{-3}{2\lambda}$

$\left(\dfrac{5}{2\lambda}\right)^2 + \left(\dfrac{-3}{2\lambda}\right)^2 = 136$

$\dfrac{25}{4\lambda^2} + \dfrac{9}{4\lambda^2} = 136$

$\dfrac{34}{4\lambda^2} = 136$

$\dfrac{17}{2\lambda^2} = 136$

$17 = 272\lambda^2$

$\dfrac{17}{272} = \lambda^2$

$\dfrac{1}{16} = \lambda^2$

$\lambda = \pm \frac{1}{4}$

When $\lambda = \frac{1}{4}$:
$\rightarrow x = 10$
$\rightarrow y = -6$

When $\lambda = -\frac{1}{4}$:
$\rightarrow x = -10$
$\rightarrow y = 6$

$f(10,-6) = 68$ Max
$f(-10,6) = -68$ Min

2. Find the max and min of $f(x,y,z) = x+z$ subject to the constraint $x^2+y^2+z^2=1$.

Soln:
$1 = 2\lambda x$
$0 = 2\lambda y$
$1 = 2\lambda z$
$x^2+y^2+z^2=1$

$x = \frac{1}{2\lambda}$, $y=0$, $z = \frac{1}{2\lambda}$

$$\left(\frac{1}{2\lambda}\right)^2 + (0)^2 + \left(\frac{1}{2\lambda}\right)^2 = 1$$

$$\frac{2}{4\lambda^2} = 1$$

$$\frac{1}{2\lambda^2} = 1$$

$\lambda^2 = \frac{1}{2}$
$\lambda = \pm \frac{1}{\sqrt{2}}$

$x = \pm\frac{\sqrt{2}}{2} = \pm\frac{1}{\sqrt{2}} = z, \; y = 0$

$(-\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ and $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ are the 2 points.

$f(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}$
$= \frac{2}{\sqrt{2}}$ Max

$f(-\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}) = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}$
$= -\frac{2}{\sqrt{2}}$ Min

3. Find the max and min of $f(x,y,z) = x-y+z$
subject to the constraint $x^2+y^2+z^2=2$

$1 = 2\lambda x$
$-1 = 2\lambda y$
$1 = 2\lambda z$
$x^2+y^2+z^2 = 2$

$x = \frac{1}{2\lambda}, \quad y = \frac{-1}{2\lambda}, \quad z = \frac{1}{2\lambda}$

$3\left(\frac{1}{2\lambda}\right)^2 = 2$

$\frac{1}{4\lambda^2} = \frac{2}{3}$

$\lambda^2 = 3/8$
$\lambda = \pm\sqrt{3/8}$

When $\lambda = \sqrt{\frac{3}{8}}$

$$x = \frac{1}{2\lambda}$$

$$= \frac{1}{2\sqrt{3/8}}$$

$$= \frac{1}{\sqrt{12/8}}$$

$$= \frac{1}{\sqrt{3/2}}$$

$$= \sqrt{2/3}$$

$$y = -\sqrt{2/3}$$

$$z = \sqrt{2/3}$$

When $\lambda = -\sqrt{3/8}$

$$x = -\sqrt{2/3}$$

$$y = \sqrt{2/3}$$

$$z = -\sqrt{2/3}$$

$f(\sqrt{2/3}, -\sqrt{2/3}, \sqrt{2/3}) = \sqrt{2/3}$ Max

$f(-\sqrt{2/3}, \sqrt{2/3}, -\sqrt{2/3}) = -\sqrt{2/3}$ Min