Dr. Tony Diana

DATA 602

Calculus Cheat Sheet

Gradient

Gradient for functions with respect to a real-valued matrix **A** is defined as the matrix of partial derivatives of **A** and is denoted as follows:

$$\nabla_A f(A) \in \mathbb{R}^{m\times n} = \begin{bmatrix} \frac{\partial f(A)}{\partial A_{11}} & \frac{\partial f(A)}{\partial A_{12}} & \cdots & \frac{\partial f(A)}{\partial A_{1n}} \\ \frac{\partial f(A)}{\partial A_{21}} & \frac{\partial f(A)}{\partial A_{22}} & \cdots & \frac{\partial f(A)}{\partial A_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(A)}{\partial A_{m1}} & \frac{\partial f(A)}{\partial A_{m2}} & \cdots & \frac{\partial f(A)}{\partial A_{mn}} \end{bmatrix}$$

$$\left(\nabla_A f(A)\right) ij = \frac{\partial f(A)}{\partial A_{ij}}$$

In vector calculus, the gradient is a multi-variable generalization of the derivative. Whereas the ordinary derivative of a function of a single variable is a scalar-valued function, the gradient of a function of several variables is a vector-valued function. The gradient stores all the partial derivative information of a multivariable function.

Hessian

Gradient is the first derivative for functions of vectors, whereas hessian is the second derivative. We will go through the notation now:

$$\nabla_x^2 f(x) \in \mathbb{R}^{n\times m} = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

Similar to the gradient, the hessian is defined only when f(x) is real-valued.

The Hessian Matrix is a square matrix of second ordered partial derivatives of a scalar function. It is of immense use in linear algebra as well as for determining points of local maxima or minima.

The Hessian matrix or Hessian describes the local curvature of a function of many variables. Hesse originally used the term "functional determinants."

Determinant

Determinant shows us information about the matrix that is helpful in linear equations and also helps in finding the inverse of a matrix.

For a given matrix **X**, the determinant is shown as follows:

$$X = \begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}$$

$$\det(X) = a(ei - fh) - b(di - fg) - c(dh - eg)$$

In linear algebra, the **determinant** is a scalar value that can be computed from the elements of a square matrix and encodes certain properties of the linear transformation described by the matrix. The **determinant** of a matrix A is denoted det(A), det A, or |A|.