Matrix Calculus

Page Student Notebooks

1. Conadient

Suppose that f: Rmxn -> R is a function that

takes as imput a matrix A of size mxn and

neturns a neal value.

is the matrix of partial derivatives, defined as:

Vaf(A) ERman =	Sf(A) Sf(A) Sf(A) SAm
o or the that the	Tall Tall I was so !!
	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
and the contraction	Sf(A) Sf(A) Sf(A)
1 14 4 5	8Ami SAmz 8Amn
FOR THE STATE OF T	

091

$$(\nabla_A f(A))_{ij} = \frac{Sf(A)}{SA_{ii}}$$

=> If in particular, A is just a vector oce RM

(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	8+(21)
	82,
$\nabla_{x} f(x) =$	8f(7l)
	5×2
	_
	8f(x)
	SXn

=> 9t is very impartant to numerous that the gradient of a function is only defined if the function is only defined if the function is oracl-values, that is, if it notems a scalar value.

Vx (af61) + bg(a) = a Vocf(21) + b Vocg(2)

2. The Hessian

Juppose that f: R" > R is a function that takes a vactor in R" and returns a neal number.

=> Them the Hessian matrix with prespect to \$\implies \text{Consitten } \forall f(x) on Simply H is the nxn matrix of partial derivatives

- Company				
77111	8f60	8 f(x)	- ~	Sia
Trans	Sx.	8x, 8x2		5x,5xn
Vxf(x) E Rnxn=	$8^2 f(x)$	8 ² f(01)		8 ² f(01)
	82282	SX2		8×28×2
ne notes have a		It was		2_ }
principal de la constant de la const	82F(x)	84fby.		8 f(x)
No.	Exnex,	Sx, 8x,		8×2

00

$$\left(\nabla_{\infty}^{2}f(\pi)\right)_{i;}=\frac{S^{2}f(\pi)}{S\pi;S\pi;}$$

=> Hessian is clarge symmetric, since

 $\frac{S^{2}f(x)}{S^{2}(x)} = \frac{S^{2}f(x)}{S^{2}(x)}$

Similar to the gradient, the Hessian is defined only when for is need-valued.

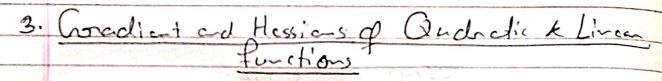
> Chadient of Gradient is not hessian for a vector volved function.

The gradient of the function is a vector and we cannot take the gradient of a vector.

=> 9f we don't mind being a little bit sloppy we can say that

 $\nabla_{\alpha}^{2} f(\alpha) = \nabla_{\alpha} (\nabla_{\alpha} f(\alpha))^{T}$

-> As long as we understand that this needly mans taking the gradient of each entry of (Voif (51)), not the gradient of the cubile vactor.



i)
$$\nabla_{x} b x = b'$$

2)
$$\nabla_X \propto^T A > L = (A + A^T) \propto \rightarrow 2A \propto Lif A symmetry$$

3)
$$\nabla_{x} x^{T} A x = (A + A^{T}) \longrightarrow 2A [if A symmetri)$$

4)
$$\nabla_{A} |A| = [ads(A)J = |A|A^{-T}]$$

5)
$$\nabla_{A} \log |A| = \frac{1}{|A|} \nabla_{A} |A| = A^{-1}$$