50.039 Theory and Practice of Deep Learning
Theory Homework 2

Joel Huang 1002530

February 13, 2019

1 Linear Algebra

Matrix inverse

$$XA + A^\top = I$$

$$XA = I - A^\top$$

Taking the inverse of A on both sides,

$$XAA^{-1} = (I - A^\top)A^{-1}$$

$$XI = IA^{-1} - A^T A^{-1}$$

$$X = (1 - A^\top)A^{-1}$$

Matrix inverse and transpose

$C - 2A^\top$ is invertible.

$$X^\top C = [2A(X + B)]^\top$$

$$X^\top C = 2A^\top X^\top + 2A^\top B^\top$$

$$X^\top C - 2A^\top X^\top = 2A^\top B^\top$$

$$X^\top (C - 2A^\top) = 2(AB)^\top$$

$$X^\top (C - 2A^\top)(C - 2A^\top)^{-1} = 2(AB)^\top (C - 2A^\top)^{-1}$$

$$X = (2(AB)^\top (C - 2A^\top)^{-1})^\top$$

Matrix inverse

$$(Ax - y)^\top A = 0$$

$$(Ax - y)^\top = 0^\top$$

$$Ax = y$$

In order to remove A from the left side, the inverse of A must be defined, A^{-1}.

$$x = A^{-1}y$$

Positive definiteness and invertibility

Any positive definite matrix is invertible.

$$(Ax - y)^\top A = -x^\top B$$

$$x^\top A^\top A - y^\top A = -x^\top B$$

$$x^\top A^\top A + x^\top B = y^\top A$$

$$x^\top (A^\top A + B) = y^\top A$$

In the case where A is invertible,

$$Ax \neq 0$$

Necessarily, $A^\top A$ is positive definite, as per the definition of a positive definite matrix:

$$x^\top A^\top A z > 0$$

$$(Ax)^\top (Ax) > 0$$

Since $A^\top A$ is positive definite, and B is positive definite, then $(A^\top A + B)$ is also positive definite, and therefore invertible.

$$x^\top (A^\top A + B)(A^\top A + B)^{-1} = y^\top A(A^\top A + B)^{-1}$$

$$x^\top = y^\top A(A^\top A + B)^{-1}$$

$$x = (y^\top A(A^\top A + B)^{-1})^\top$$

2 Directional derivative

For a function $f : \mathbb{R}^n \mapsto \mathbb{R}^1$, and some arbitrary direction v, the directional derivative $\nabla f(x) \cdot v$ can be seen as a projection along v. We need to choose the right v such that the maximum value of $\nabla f(x) \cdot v$ is achieved. Using the geometric definition of the dot product,

$$\nabla f(x) \cdot v = \|\nabla f(x)\| \, \|v\| cos(\theta)$$

where θ is the angle between the gradient and v. This function is at its maximum when $cos(\theta) = 1$, which implies $\theta = 0°$. $\nabla f(x)$ is indeed codirectional with v.