50.039 – Theory and Practice of Deep learning

Alex

Week 02

[The following notes are compiled from various sources such as textbooks, lecture materials, Web resources and are shared for academic purposes only, intended for use by students registered for a specific course. In the interest of brevity, every source is not cited. The compiler of these notes gratefully acknowledges all such sources.]

Due: week3 wednesday, 6pm

1 Task1

Let X,A be arbitrary matrices and A invertible. Solve for X:

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

Let X,A,B be arbitrary matrices and $C - 2A^{T}$ invertible. Solve for X:

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

Let $x \in RR^{n}$, $y \in \mathbb{R}^{d}$, $A \in \mathbb{R}^{d \times n}$. What term must be invertible, so that

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

can be solved for x? What is the solution?

As above, let $B \in \mathbb{R}^{n \times n}$ be positive definite, then solve for x

$$(Ax - y)^{\top}A + x^{\top}B = 0$$

2 Task2

Proof that for a function $f : \mathbb{R}^n \to \mathbb{R}^1$ differentiable in x, the gradient direction is the direction where the function locally increases fastest.

Hint:

- you can use that the direction derivative in direction v from point x is given as $\nabla f(x) \cdot v$
- consider a constrained optimization problem in argument v. You may consider a linear combination $w = \cos(a)v_1 + \sin(a)v_2$, because its L2-norm can be computed easily if v_1 and v_2 are orthogonal. Though you need to think what to use for v_1 and v_2