Optimization Algorithms
Weekly Exercise 0

Marc Toussaint
Learning & Intelligent Systems Lab, TU Berlin
Marchstr. 23, 10587 Berlin, Germany

Winter 2024/25

You do not have to prepare for the first tutorial. You should be able to solve these exercises directly. Please volunteer to solve these exercises (or help each other to solve them jointly) at the board.

1 Matrix equations

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

$$XA + A^\top = \mathbf{I} \tag{1}$$

b) Let X, A, B be arbitrary matrices, $(C - 2A^\top)$ invertible. Solve for X:

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

c) Let $x \in \mathbb{R}^n, y \in \mathbb{R}^d, A \in \mathbb{R}^{d\times n}$. A obviously *not* invertible, but let $A^\top A$ be invertible. Solve for x:

$$(Ax - y)^\top A = \mathbf{0}_n^\top \tag{3}$$

d) As above, additionally $B \in \mathbb{R}^{n\times n}$, B positive-definite. Solve for x:

$$(Ax - y)^\top A + x^\top B = \mathbf{0}_n^\top \tag{4}$$

2 Multivariate Calculus

Compute the following Jacobian tensors

a) $\frac{\partial}{\partial x} x$, where x is a vector

b) $\frac{\partial}{\partial x} x^\top A x$, where A is a matrix

c) $\frac{\partial}{\partial x} f(x)^\top A g(x)$, where f and g are vector-valued functions

d) $\frac{\partial}{\partial x} \frac{Ax}{||Ax||}$, where A is a matrix and $||\cdot||$ is the 2-norm

3 Minimization

You can minimize a function $f(\beta)$ by computing its gradient $\nabla_\beta f(\beta)$ and solving $\nabla_\beta f(\beta) = 0$ w.r.t. β.
Given a fixed $y \in \mathbb{R}^m$, $X \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}^+$, find the value of $\beta \in \mathbb{R}^n$ that minimizes:

$$||y - X\beta||^2 + \lambda||\beta||^2 \tag{5}$$

Did you find a local or global optimum? Justify your answer.

4 Projections

a) In $\mathbb{R}^n$, a plane (through the origin) is described by the linear equation

$$c^\top x = 0 \ , \tag{6}$$

where $c \in \mathbb{R}^n$ parameterizes the plane. Provide the matrix that describes the orthogonal projection of a point $x_0 \in \mathbb{R}^n$ onto this plane. (Hint: the solution has the form of 'Identity matrix minus a rank-1 matrix').

b) In $\mathbb{R}^n$, we have k linearly independent vectors $\{v_i\}_{i=1}^k$, which form the matrix $V = (v_1, .., v_k) \in \mathbb{R}^{n \times k}$. We want to derive the projection of $x_0 \in \mathbb{R}^n$ into the subspace generated by the columns of V as an optimization problem.

Any point in the subspace can be expressed as $V\alpha,\ \alpha \in \mathbb{R}^k$. Therefore, the projection of x_0 into V is $V\alpha^*$, where

$$\alpha^*(x_0) = \underset{\alpha \in \mathbb{R}^k}{\operatorname{argmin}} \|x_0 - V\alpha\|^2 \ . \tag{7}$$

Derive the expression for the optimal $\alpha^*(x_0)$ from the optimality principles.