Homework 0

Lionel Tondji
Warm-up

July 13, 2024

Exercise 1: An $n \times n$ symmetric real matrix M is said to be positive-definite (p.d) if

$$x^\top Mx > 0, \forall\ 0 \neq x \in \mathbb{R}^n.$$

It is said to be positive-semidefinite (p.s.d) if

$$x^\top Mx \geq 0, \forall\ x \in \mathbb{R}^n$$

1. Show that if M is p.s.d and $\varepsilon > 0$, then $M + \varepsilon I$ is p.d.
2. Deduce that $M^\top M + \varepsilon I$ is p.d for any matrix M and any $\varepsilon > 0$
3. Let $\mathrm{Null}(M) = \{x \in \mathbb{R}^n : Mx = 0\}$. Show that if $\mathrm{Null}(M) = \{0\}$ then $M^\top M$ is p.d.

Exercise 2: Let Ω be a finite set and $X : \Omega \mapsto \mathbb{R}, w \in \Omega \mapsto X(w)$ be a real valued function. We further suppose that X has a discrete number of output given by $X(\Omega) = \{x_1, \ldots, x_n\}$. We refer to this X as a random variable (rv). We define its expectation as

$$E(X) = \sum_{i=1}^{n} x_i P[X = x_i]$$

. Let Y be another r.v. We define the conditional expectation of X on $Y = y_j$ as

$$E[X/Y = y_j] = \sum_{i=1}^{n} x_i P[X = x_i/y = y_j]$$

. It is the expected value of X but given that $Y = y_j$ has occured.

1. Show that $E[E[X/Y]] = E[X]$
2. Let $V = (V_1, \ldots, V_n)^\top, W = (W_1, \ldots, W_n)^\top$ be two independent r.v. Show that

$$E[\langle V, W\rangle/W] = \langle E[V], W\rangle$$

Exercise 3: All the algorithm we will discuss in the course generate a sequence of random vectors x_t that converge to a desired x^* in some sense. In particular, we focus on two forms of convergence, either showing that the difference of function values converges

$$f(x_t) - f^* \to 0$$

or norm difference of the iterates converges

$$\|x_t - x^*\| \to 0$$

Two important questions:

- How fast is this convergence ?
- given an ε accuracy how many iterations t are needed before $f(x_t) - f^* < \varepsilon$ or $\|x_t - x^*\| < \varepsilon$.

1. Consider a sequence $(u_t)_t \in \mathbb{R}_+$ of positive scalars that converge to zero according to

$$u_t \leq \frac{C}{t}$$

where $C > 0$. Given $\varepsilon > 0$, show that

$$t \geq \frac{C}{\varepsilon} \Rightarrow u_t < \varepsilon$$

We refer to this result as a $\mathcal{O}(1/\varepsilon)$ iteration complexity. This type of convergence is known as sublinear convergence.

2. Consider a sequence $(u_t)_t \in \mathbb{R}_+$ of positive scalars that converge to zero according to

$$u_t \leq q^t u_0$$

where $0 \leq q < 1$. Given $1 > \varepsilon > 0$, and using the fact that

$$\frac{1}{1-q} \cdot \log(\frac{1}{q}) \geq 1,$$

show that

$$t \geq \frac{1}{1-q} \cdot \log(\frac{1}{\varepsilon}) \Rightarrow u_t < \varepsilon u_0$$

We refer to this result as a $\mathcal{O}(log(1/\varepsilon))$ iteration complexity. This type of convergence is known as linear convergence at a rate q^t.

Exercise 1: Compute ∇f and $\nabla^2 f$ for the following functions $f : \mathbb{R}^n \to \mathbb{R}$.

1. $f(x) = a^\top x$
2. $f(x) = x^\top M x$
3. $f(x) = x^\top M x + b^\top x + c$