E0 230
Computational Methods of Optimization
Tutorial 1

Nov. 7, 2020

---

1. Suppose $x,\ y \in \mathbb{R}^n$. Prove the following statements.
   (a) $|\|x\| - \|y\|| \leq \|x - y\| \leq \|x\| + \|y\|$
   (b) $2x^T y = \|x\|^2 + \|y\|^2 - \|x - y\|^2$
   (c) $\|x\|_2 \leq \sqrt{n}\|x\|_\infty$
   (d) $\|x\|_1 \leq n\|x\|_\infty$
   (e) $\|x\|_1 \leq \sqrt{n}\|x\|_2$

2. Note that $H_f(x)$ denotes the Hessian of a function $f : \mathbb{R}^n \to \mathbb{R}$
   (a) Show that $e^x \geq 1 + x$.
   (b) Suppose $n = 1$, and $f^{(k)}$, the kth derivative of f w.r.t x is absolutely continuous. Show that given
   $$f(x) = f(x_0) + f'(x_0)(x - x_0) + ... \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k + R_k$$
   we have
   $$R_k = \int_{x_0}^{x} \frac{f^{(k+1)}(t)}{k!}(x - t)^k dt.$$
   (c) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable, such that $\max_{x,i} |\lambda_i(H_f(x))| = M < \infty$, where $\lambda_i(H_f(x))$ is the ith eigenvalue of $H_f(x)$. Show that there exists a constant L such that, for each $x,\ y$, we have
   $$f(y) - f(x) \leq \nabla f(x)^T (y - x) + \frac{L}{2}\|y - x\|^2$$
   What is the smallest value of L for which this expression is satisfied?

3. Suppose we have matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times n}$, $m \neq n$
   (a) Suppose A has n orthogonal eigenvectors. Show that we can write $A = V^T \Lambda V$, where the columns of V are the eigenvectors of A, and the diagonal elements of Λ are the eigenvalues of A. When are we unable to use this decomposition?
   (b) Suppose B is of rank r. What are the ranks of BB^T and $B^T B$? With this result, what can you say about the row and column ranks of B and B^T?
   (c) Let $p(l) = \det(lI - A)$. For any A, we have $p(A) = 0$ (this is the Cayley-Hamilton theorem). Show that $p(A) = p(P^{-1}AP)$ for any invertible matrix P. What does this say about the eigenvalues of A and $P^{-1}AP$?
   (d) Show that we can decompose $B = U\Sigma V^T$, where Σ is diagonal and positive-semidefinite, and U and V are unitary matrices.

(e) Show that A, B are equivalent if and only if, for all vectors $v \in \mathbb{R}^n$, $Av = Bv$.

(f) The Frobenius norm of a matrix is given by $\|B\|_F = \sqrt{\mathrm{Tr}(B^T B)}$. Show that $\|B\|_F = \sqrt{\sum_k \sigma_k(B)^2}$, where $\sigma_k(B)$ is the kth largest singular value of B.

(g) Show that $B_k = \sum_{i=1}^k \sigma_i u_i v_i^T$, where $k \leq \min\{m, n\}$, satisfies $\|B - B_k\|_F \leq \|B - C\|_F$ for any $C \in \mathbb{R}^{m \times n}$ of rank k.

(h) Consider the function $f : U \to \mathbb{R}$, where $f(x) = x^T Ax$ and U is the set of all unit vectors of dimension n. Show that range$(f) \subseteq [\lambda_{min}(A), \lambda_{max}(A)]$.

(i) What is the solution to

$$s^* = \operatorname*{argmax}_x \frac{\|Bx\|_2}{\|x\|_2}?$$

*R*emark: s^* is the Spectral norm of B, denoted by $\|B\|_2$.

4. Consider the polynomial

$$p(x, y, z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2.$$

Show that

$$f^* = \inf_{x,y,z} p(x, y, z) = 0.$$

5. Suppose A, B are symmetric and that the problems

$$\text{(P1)} \quad \operatorname*{argmin}_x x^T Ax \quad \text{and} \quad \text{(P2)} \quad \operatorname*{argmin}_x x^T Bx$$

have unique solutions $x_{P1} = x_{P2} = 0$. What is the solution to

$$\operatorname*{argmin}_x x^T ABx,$$

and is it unique? (hint: every PD matrix has a unique, positive definite square root - can you prove this?)

6. Suppose we have m scalar data points $\{x_i\}_{i=1}^m$. What is the solution to

$$z_2 = \operatorname*{argmin}_z \sum_i (x_i - z)^2.$$

7. Suppose we have m pairs of data points (x_i, y_i), where $x_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$. Solve

$$x_* = \operatorname*{argmin}_w \sum_i (y_i - w^T x_i)^2.$$

What if there are only $r < \min\{m, n\}$ linearly independent data points? Will you still have a unique solution (show why or why not).

8. Suppose we have positive definite $A \in \mathbb{R}^{n \times n}$, and linearly independent vectors $\{v_i\}_{i=1}^m$, where $m < n$. How would you convert the problem

$$\operatorname*{argmin}_x x^T Ax \;\text{ such that } x \in \mathrm{span}\,(v_1, ..., v_m)$$

into an unconstrained problem? Does this problem have a unique solution? If so, under what conditions would this problem not have a unique solution?