1 Fundamentals of unconstrained optimization

1. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, write A as a sum of rank 1 matrices. Use the Eigenvalue decomposition of A.

2. Let $a \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Define $f_1(x) = a^\top x$ and $f_2(x) = \frac{1}{2}x^\top Ax$. Compute ∇f_1, $\nabla^2 f_1$, ∇f_2, and $\nabla^2 f_2$.

3. Let f be a strictly convex function. Show that if x^* is a local minimizer then it is a global minimizer.

4. Show that $-\nabla f$ is a descent direction.

5. Let $f(x, y) = (x + y^2)^2$. Show that $p = [-1, 1]^\top$ is a descent direction at $[1, 0]^\top$. State the steepest descent direction and Newton's direction. What is the optimum of this function? For Newton's method how many steps do we need to reach such optimum?

2 Newton and Quasi-Newton methods

1. State Newton's update formula.

2. State two differences between Newton's method and steepest descent (other than the formulas).

3. Let $\nabla^2 f(x) \in \mathbb{R}^{n\times n}$ be a constant matrix (i.e. independent of x). Let $d \in \mathbb{R}^n$ such that it solves the following system:

$$0 = \nabla f(y) + \nabla^2 f(y) d.$$

   Prove that $y + d$ is a stationary point of f.

4. Let:

$$B_k = (I - \rho\gamma s^\top) B_{k-1} (I - \rho s \gamma^\top) + \rho\gamma\gamma^\top,$$

   with $\rho = 1/(\gamma^\top s)$, γ, s fixed vectors.

   (a) Which update formula is this?

   (b) What does B_k approximates? What is the difference between this method and Newton's method?

   (c) Prove that $B_k s = \gamma$

   (d) Prove that if B_{k-1} is spd and $s^\top\gamma > 0$ then B_k is also spd. What is the practical importance of this?

   (e) Prove that if $B_{k-1} s = \gamma$ then $B_k = B_{k-1}$. What does this mean?

3 KKT Conditions

1. State the KKT conditions for the following optimization problem:

$$
\begin{aligned}
\min\ & f_0(x) & & \\
\text{s.t. }\ & f_k(x) = 0 & & k = 1, ..., m \\
& g_j(x) \leq 0 & & j = 1, ..., r
\end{aligned}
$$

2. Solve the following optimization problem using the KKT conditions:

$$
\begin{aligned}
\min\ & -xy \\
\text{s.t. }\ & x + y = 10
\end{aligned}
$$

3. Using the KKT conditions, find the point on the circle $x^2 + y^2 = 80$ that is closest to $(1, 2)$. You are going to find to KKT points, what is the relation between that other non optimal point and $(1, 2)$?

4 Duality

1. What is strong duality? What is weak duality?
2. State the definition of the Lagrange dual function
3. How does the Lagrange dual function relate to the dual LP problem?
4. Find the Lagrange dual function of the following problem:

$$\begin{aligned} \min\ & x^\top x \\ \text{s.t. } & Ax = b \end{aligned}$$

5. Consider the following LP in standard form:

$$\begin{aligned} \min\ & c^\top x \\ \text{s.t. } & Ax = b \\ & x \geq 0 \end{aligned}$$

   Compute the Lagrangua dual function. What can you tell about the dual LP problem from this function?

5 Quadratic programming

1. Consider the following optimization problem:

$$\min q(x) = \frac{1}{2}x^\top Qx + c^\top x \tag{1}$$
$$\text{s.t. } Ax = 0, \tag{2}$$

   where $Q \in \mathbb{R}^{n\times n}$ is spd, $A \in \mathbb{R}^{m\times n}$ with $rank(A) = m$.

   (a) Let x^* be the optimum, prove that $q(x^*) \le 0$.

   (b) Prove that $q(x^*) = 0 \iff x^* = 0$.

2. Consider the problem of finding the point on a hyperplane H that has the minimum distance to a fixed point x_0. This hyperplane is defined as:

$$H = \{x \in \mathbb{R}^n : Ax = b\},$$

   where x_0, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ are fixed and $rank(A) = m$.

   (a) Write this as a constrained optimization problem.

   (b) Write down the Lagrange function for this problem

   (c) Deduce that the solution is given by:

$$x^* = x_0 + A^\top(AA^\top)(b - Ax_0)$$
$$\lambda^* = -(AA^\top)^{-1}(b - Ax_0).$$