Friederike Horn & Bileam Scheuvens

Justify all your claims.

Exercise 1 (Netwon Method vs. Gradient Descent)

Solution

a)

$$\nabla \frac{1}{2} x^t Q x - b^t x = Qx - b \stackrel{!}{=} 0$$
$$\Rightarrow Qx = b$$

Divide by Q, given that Q is symmetric and PD and therefore invertible.

$$x = Q^{-1} b$$

This is a minimum and therefore the optimal solution to the problem, if the Hessian is positive definite.

$$\nabla^2 \frac{1}{2} x^t Q x - b^t x = Q$$

Since we know Q to be PD, $x^* = Q^{-1} b$.

b)

$$||x_{k+1} - x^*|| = ||x_k - \alpha \nabla f(x_k) - x^*||$$
$$= ||(x_k - x^*) - \alpha(Qx_k + b)||$$

Since $b = Qx^*$

$$= ||(x_k - x^*) - \alpha Q(x_k + x^*)||$$
$$= ||(x_k - x^*)(I - \alpha Q)||$$
$$\leq ||(x_k - x^*)|| \cdot ||(I - \alpha Q)||$$

Since the largest stretch that $1 - Q$ can apply, is none - the smallest stretch by Q:

$$= ||(x_k - x^*)|| \cdot (1 - \alpha \lambda_{min})$$

Let $\alpha = \frac{2}{\lambda_{max} + \lambda_{min}}$

$$= ||(x_k - x^*)|| \cdot (1 - \frac{2\lambda_{min}}{\lambda_{max} + \lambda_{min}})$$
$$= ||(x_k - x^*)|| \cdot (\frac{\lambda_{max} + \lambda_{min}}{\lambda_{max} + \lambda_{min}} - \frac{2\lambda_{min}}{\lambda_{max} + \lambda_{min}})$$
$$= ||(x_k - x^*)|| \cdot \left(\frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} \right)$$

c)

$$x_{k+1} = x_k - Q^{-1} \nabla (\frac{1}{2} x_k^t Q x - b^t x_k)$$
$$x_{k+1} = x_k - Q^{-1}(Qx_k - b)$$
$$x_{k+1} = x_k - Q^{-1} Q x_k + Q^{-1} b$$
$$x_{k+1} = x_k - I x_k + Q^{-1} b$$
$$x_{k+1} = Q^{-1} b$$
$$x^* = Q^{-1} b$$

Exercise 2 (Convexity and Continuity)

Solution

a) Imagine y to be any point between x and z, analogous to how continuity is defined in general. We know that:

$$f(tx + (1-t)z) < tf(x) + (1-t)f(z)$$

Let $y = tx + (1-t)z$. Then:

$$f(y) < tf(x) + (1-t)f(z)$$

Subtract $f(x)$

$$f(y) - f(x) < (t-1)f(x) + (1-t)f(z)$$

$$f(y) - f(x) < (t-1)(f(x) - f(z))$$

Divide by $y - x$, since $x < y$ and therefore $x - y \neq 0$

$$\frac{f(y) - f(x)}{y - x} < \frac{(t-1)(f(x) - f(z))}{y - x}$$

Choose $t = \frac{y-x}{z-x} + 1$

$$\frac{f(y) - f(x)}{y - x} < \frac{f(x) - f(z)}{z - x}$$

The argument is analogous for the slope between $z - x$ to $z - y$, concluding the proof.

b)

c) f is not necessarily continous in a and b as exemplified by:

$$f(x) = \begin{cases} x^2 & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

Exercise 3 (Convexity)

Solution

a)

b) Consider the logarithm. log is concave, since its second derivative $-\frac{1}{x^2}$ is negative. Since the logarithm is monotone it preserves inequality and we can instead consider

$$log(\frac{1}{n}\sum_{i=1}^{n} x_i) \leq log((\prod_{i=1}^{n} x_i)^{\frac{1}{n}})$$

$$log(\frac{1}{n}\sum_{i=1}^{n} x_i) \leq \frac{1}{n}log(\prod_{i=1}^{n} x_i)$$

$$log(\frac{1}{n}\sum_{i=1}^{n} x_i) \leq \frac{1}{n}\sum_{i=1}^{n} log(x_i)$$

This holds by the result from a).

Exercise 4 (Dual Problem)

Solution

a)

$$f(x,y) = max \;\; x + y$$
$$\text{subject to } x^2 + 2y^2 \leq 5$$

Bring to normal form by minimizing and bringing constraint to the form ≤ 0

$$f(x) = min \;\; -x - y$$
$$\text{subject to } x^2 + 2y^2 - 5 \leq 0$$
$$\mathcal{L}(x,y) = min \;\; -x - y - \lambda(x^2 + 2y^2 - 5)$$

Solve for stationarity:

$$\nabla\mathcal{L} \overset{!}{=} 0$$
$$\frac{\partial L}{\partial x} = 1 + 2\lambda x$$
$$\Rightarrow x = \frac{1}{2\lambda}$$
$$\frac{\partial L}{\partial y} = 1 + 4\lambda y$$
$$\Rightarrow y = \frac{1}{4\lambda}$$

Solve for primal feasibility:

$$x^2 + 2y^2 - 5 \leq 0$$
$$\frac{1}{4\lambda^2} + 2\frac{1}{16\lambda^2} - 5 \leq 0$$
$$\frac{1}{4\lambda^2} + \frac{1}{8\lambda^2} - 5 \leq 0$$
$$\frac{3}{8\lambda^2} \leq 5$$
$$\frac{3}{8} \leq 5\lambda^2$$
$$\frac{3}{40} \leq \lambda^2$$
$$\sqrt{\frac{3}{40}} \leq \lambda$$
$$\Rightarrow x = \frac{1}{2\sqrt{\frac{3}{40}}} = \frac{1}{\sqrt{\frac{4 \cdot 3}{40}}} = \sqrt{\frac{10}{3}}$$
$$\Rightarrow y = \sqrt{\frac{5}{3}}$$

The constraint is active since $\lambda > 0$ (dual feasibility). This makes intuitive sense since the function would otherwise be unbounded and the optimal solution would be $x = y = \infty$.

b) The Lagranian of the problem is:

$$L(x,\lambda) = min \;\; -c^t x - \lambda(Ax - b)$$

Therefore we obtain the dual function $g(\lambda) = \underset{x}{inf}\,\mathcal{L}(x,\lambda)$.