MATH520 Exam 1

Fernando

March 1, 2024

Problem 1

Part a

There are no feasible directions. This is easy to see geometrically because Ω is a circle so any direction will take us out of the circle. Now let's prove this formally by contradiction. Suppose that d is a feasible direction at x^*, then there exits $\alpha_0 > 0$ such that $x^* + \alpha d \in \Omega$ for all $\alpha \in [0, \alpha_0]$. Notice that

$$\begin{aligned}
||x^* + \alpha d||^2 &= 1 \\
||x^*||^2 + \alpha^2 ||d||^2 + 2\alpha(d \cdot x^*) &= 1 \\
\alpha^2 ||d||^2 + 2\alpha(d \cdot x^*) &= 0,
\end{aligned}$$

which is a quadratic equation on α so it has at most 2 solutions, which contradicts the fact that α can be any value on $[0, \alpha_0]$.

Part b

The FONC condition is trivially satisfied by every point in Ω, because in part a we saw that there are no feasible directions.

Part c

It is not useful, it doesn't eliminate any point.

Part d

If we define

$$g(\theta) := f(\cos(\theta), \sin(\theta))$$

then by the chain rule we get

$$\begin{aligned}
g'(\theta) &= Df(\cos(\theta), \sin(\theta)) \cdot [-\sin(\theta), \cos(\theta)]^T \\
&= [-\sin(\theta), \cos(\theta)] \cdot \nabla f(\cos(\theta), \sin(\theta))
\end{aligned}$$

Then if $x^* \in \Omega$ is a local minimizer there is a θ^* such that $[\cos(\theta^*), \sin(\theta^*)] = x^*$ and $g'(\theta^*) = 0$ (unconstrained FONC).

So

$$\begin{aligned} 0 = g'(\theta^*) &= [-\sin(\theta^*), \cos(\theta^*)] \cdot \nabla f(\cos(\theta^*), \sin(\theta^*)) \\ &= [-\sin(\theta^*), \cos(\theta^*)] \cdot \nabla f(x^*). \end{aligned}$$

Notice that $[\cos(\theta^*), \sin(\theta^*)] \cdot [-\sin(\theta^*), \cos(\theta^*)]^T = 0$, in other words: $[-\sin(\theta^*), \cos(\theta^*)]^T$ is a perpendicular vector to x^*.

So we can say that if $x^* \in \Omega$ is a local minimizer, then $d^T \nabla f(x^*) = 0$ for all d perpendicular to x^*.

Problem 2

Part a

Lemma: Under the conditions of the problem

$$\alpha_k = -\frac{d^{(k)T} g^{(k)}}{d^{(k)T} Q d^{(k)}}$$

Proof:

Let $h(\alpha) := f(x^{(k)} + \alpha_k d^{(k)})$. Then

$$h'(\alpha) = d^{(k)T} \nabla f(x^* + \alpha d^{(k)}) = d^{(k)T}(Q(x^{(k)} + \alpha d^{(k)}) - b).$$

By the definition of α_k it has to satisfy the FONC

$$h'(\alpha_k) = 0,$$

so we get the equation

$$0 = d^{(k)T}(Q(x^{(k)} + \alpha_k d^{(k)}) - b),$$

and solving for α_k we get

$$\alpha_k = \frac{d^{(k)T} b - d^{(k)T} Q x^{(k)}}{d^{(k)T} Q d^{(k)}} = -\frac{d^{(k)T} g^{(k)}}{d^{(k)T} Q d^{(k)}},$$

which concludes the proof of the lemma.

Now let's do part a.

$$\begin{aligned} 2w_k &= 2(V(x^{(k)}) - V(x^{(k+1)})) \\ &= 2(V(x^{(k)}) - V(x^{(k+1)})) \\ &= (x^{(k)} - x^*)^T Q (x^{(k)} - x^*) - (x^{(k+1)} - x^*)^T Q (x^{(k+1)} - x^*) \\ &= x^{(k)T} Q x^{(k)} - x^{(k)T} Q x^* - x^{*T} Q x^{(k)} + x^{*T} Q x^* \\ &\quad - x^{(k+1)T} Q x^{(k+1)} + x^{(k+1)T} Q x^* + x^{*T} Q x^{(k+1)} - x^{*T} Q x^* \end{aligned}$$

replacing $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$ and simplifying we get that

$$\begin{aligned}
2w_k &= \alpha_k d^{(k)T} Q x^* + \alpha_k x^{*T} Q d^{(k)} - \alpha_k d^{(k)T} Q x^{(k)} - \alpha_k x^{(k)T} Q d^{(k)} \\
&\quad - \alpha_k^2 d^{(k)T} Q d^{(k)} \\
&= 2\alpha_k d^{(k)T} Q x^* - 2\alpha_k d^{(k)T} Q x^{(k)} - \alpha_k^2 d^{(k)T} Q d^{(k)} \\
&= -2\alpha_k d^{(k)T} g^{(k)} - \alpha_k^2 d^{(k)T} Q d^{(k)} \\
&= -2\alpha_k d^{(k)T} g^{(k)} - \alpha_k^2 d^{(k)T} Q d^{(k)} \\
&= \frac{(d^{(k)T} g^{(k)})^2}{d^{(k)T} Q d^{(k)}} \text{ (by the previous lemma).}
\end{aligned}$$

Part b

Suppose that $x^{(k)} \neq x^*$ and consider

$$r_k = \frac{V(x^{(k)}) - V(x^{(k+1)})}{V(x^{(k)})} = \frac{\frac{\left(d^{(k)T} g^{(k)}\right)^2}{d^{(k)T} Q d^{(k)}}}{(x^{(k)} - x^*)^T Q (x^{(k)} - x^*)},$$

notice that $r_k \geq 0$ because Q is positive definite, also $r_k \leq 1$ because $r_k = 1 - V(x^{(k+1)})/V(x^{(k)})$ and $V \geq 0$, then

$$V(x^{(k+1)}) = (1 - r_k) V(x^{(k)}).$$

For the case $x^{(k)} = x^*$ the equation above is satisfied for any r_k because $V(x^{(k)}) = V(x^{(k+1)}) = 0$, so let's pick $r_k = 1$ for this case (we could have picked any other number with norm at most one). Now we have defined r_k properly and we can write

$$V(x^{(k+1)}) = \prod_{i=0}^{k} (1 - r_i) V(x^{(0)}) = V(x^{(0)}) \prod_{i=0}^{k} (1 - r_i).$$

In order to analyse $\lim_{k\to\infty} V(x^{(k+1)})$ we consider two cases.

Case 1: There is an infinite number of positive r_i's. In this case the limit

$$\lim_{k\to\infty} V(x^{(k+1)}) = V(x^{(0)}) \lim_{k\to\infty} \prod_{i=0}^{k} (1 - r_i)$$

is convergent because we have a decreasing sequence bounded below (by 0).

Case 2: There is a finite number of positive r_i's. In this case at some point all the r_i's become 0 and we are left with a tail of ones in the product which again converges.

Conclusion: $\lim_{k\to\infty} V(x^{(k+1)})$ is convergent.

Now let's finish part b

From part a we get

$$V(x^{(k+1)}) = V(x^{(0)}) - \sum_{i=0}^{k} w_i,$$

so

$$\sum_{i=0}^{k} w_i = V(x^{(0)}) - V(x^{(k+1)})$$

taking the limit as $k \to \infty$ and using the fact that $\lim_{k\to\infty} V(x^{(k+1)})$ is convergent we get the result.

Part c

From the formula given we have

$$||g^{(k)}||^2 \cos^2(\theta_k) = \frac{(d^{(k)T} g^{(k)})^2}{||d^{(k)}||^2}$$

and by Rayleigh's inequality we get

$$||g^{(k)}||^2 \cos^2(\theta_k) = \frac{(d^{(k)T} g^{(k)})^2}{||d^{(k)}||^2} \leq 2\lambda_{\max}(Q) w_k,$$

where w_k was obtained in part a.

Taking the sum on both sides and using part b yields the result.

Part d

$$\sum_{k=0}^{\infty} \delta^2 ||g^{(k)}||^2 \leq \sum_{k=0}^{\infty} \cos^2(\theta_k) ||g^{(k)}||^2 < \infty,$$

which implies that $||g^{(k)}|| \to 0$, then

$$||x^{(k)} - x^*|| = ||Q^{-1}(Qx^{(k)} - b)|| \leq ||Q^{-1}|| \cdot ||g^{(k)}||$$

Taking the limit as $k \to \infty$ we obtain the result.

Problem 3

We proceed by induction as suggested by the hint.

For $k = 0$ it is clear that $x^{(0)} = 0 \in \{0\} = \mathcal{V}_0$ and $d^{(0)} = -g^{(0)} = b \in \text{span}[b] = \mathcal{V}_1$.

For the induction step we assume that the result is true for k and we have to prove that $x^{(k+1)} \in \mathcal{V}_{k+1}$ and $d^{(k+1)} \in \mathcal{V}_{k+2}$.

Since

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$$

and by I.H. $x^{(k)} \in \mathcal{V}_k$ and $d^{(k)} \in \mathcal{V}_{k+1}$ then $x^{(k+1)} \in \mathcal{V}_{k+1}$ (Clearly $\mathcal{V}_k \subset \mathcal{V}_{k+1}$).

As for $d^{(k+1)}$ we have

$$d^{(k+1)} = -g^{(k+1)} + \beta_k d^{(k)},$$

since $d^{(k)} \in \mathcal{V}_{k+1}$ by I.H. it is enough to prove that $g^{(k+1)} \in \mathcal{V}_{k+2}$ (because $\mathcal{V}_{k+1} \subset \mathcal{V}_{k+2}$). To do this notice that $x^{(k+1)}$ is a linear combination of the vectors $b, Qb, \ldots, Q^k b$, then

$$\begin{aligned} g^{(k+1)} = Qx^{(k+1)} - b = Q\left(\sum_{i=0}^{k} c_i Q^i b\right) - b \\ = \left(\sum_{i=0}^{k} c_i Q^{i+1} b\right) - b \in \operatorname{span}[b, Qb, \ldots, Q^{k+1}b] = \mathcal{V}_{k+2}, \end{aligned}$$

which concludes the proof.