Numerical Optimisation: Assignment 8

Student Number:13064947

March 2018

1 Exercise 1

Consider the following minimisation, with $G \in R^{nxn}$, positive semi-definite, $x \in R^n$, $A \in R^{mxn}$, $c \in R^n$, $b \in R^m$.

$$\min_x f(x) = \frac{1}{2}x^T Gx + c^T x, \text{ subj to } Ax \leq b$$

1.1 State the KKT conditions for this problem

- $Ax \leq b \implies Ax - b \leq 0$, or $(Ax - b)_i \leq 0$, $\forall i \in 1, \ldots, m$
- $\nabla_x L = 0$, now $L(x, \lambda) = \frac{1}{2}x^T Gx + c^T x + \lambda^T (Ax - b)$, so $\nabla_x L = 0 \implies x = -G^{-1}(A^T \lambda + c)$
- $\lambda \geq 0$, strictly speaking $\lambda_i \geq 0$, $\forall i \in 1, \ldots, m$
- $\lambda_i (Ax - b)_i = 0$, $\forall i \in 1, \ldots, m$

The last condition kind of comes out of the blue given the inequality constraint, but I believe part b) actually cleanly addresses where this comes from.

1.2 Rewrite the problem using a vector of slack variables $y \in R^m, y \geq 0$ and give the corresponding KKT conditions

The obvious choice here is to add the following constraints, $y_i \geq 0, \forall i \in 1, \ldots, m$, $(Ax - b)_i + y_i = 0$, and then we have an equality constraint and an inequality constraint and adjust our lagrangian to include our y's. In this case we will wind up with a lagrangian in variables x, y, λ, ν.

Differences would basically be $(Ax - b)_i + y_i = 0$, $\lambda_i \geq 0, \nu_i \geq 0, y_i \geq 0$ and $L(x, y, \lambda, \mu) = \frac{1}{2}x^T Gx + c^T x + \lambda^T (Ax - b + y) - \mu^T y$ etc.

However to me at least, it seems simpler to realise that we lose no generality by adding y_i^2 instead. In this case, the slack values will be greater than or equal to zero anyway (so we don't need the positivity constraint on y_i) and we can still do our equality constraint. In this case the new KKT constraints will be.

- $(Ax-b)_i + y_i^2 = 0$, $\forall i \in 1,\ldots,m$
- $\nabla_x L = 0,$, here $L(x,y,\lambda) = \frac{1}{2}x^T Gx + c^T x + \lambda^T(Ax - b + y^2)$, so $\nabla_x L = 0 \implies x = -G^{-1}(A^T\lambda + c)$, as before.
- $\lambda_i \geq 0$, $\forall i \in 1,\ldots,m$
- $\nabla_y L = 0, \implies 2\lambda_i y_i = 0, \implies \lambda_i y_i = 0$, $\forall i \in 1,\ldots,m$, this is new and needed because of our added y_i^2's.

The last condition from our equality constraints is interesting and it shows whether the slack variable is being applied and I believe corresponds to the last constraint in exercise 1 part a) for the inequality constraint.

To show this consider that our last condition forces either $\lambda_i = 0$ or $y_i = 0$, but if $y_i = 0$ this means that $(Ax-b)_i = 0$ (from the first condition above). This means either $\lambda_i = 0$ or $(Ax-b)_i = 0$, or in other words that $\lambda_i(Ax-b)_i = 0$, $\forall i \in 1,\ldots,m$. But this is precisely the first condition from part a), (but using y_i^2 for our slack variable). Also the conditions $(Ax-b)_i + y_i^2 = 0$, $\forall i \in 1,\ldots,m$ and $(Ax-b)_i \leq 0$, $\forall i \in 1,\ldots,m$ are completely equivalent. So using this we can effectively derive the earlier KKT inequality conditions.

Thus I believe that in a funny way part b) is the natural precursor to part a) in these exercises - ie we start with KKT for equality constraints, look at an example of an inequality constraint - add on slack to make it an equality constraint (in this case squared slack) and derive the inequality conditions for 1a). Given this complete equivalence I hope it is ok to use the simpler form for the duality part to follow.

1.3 Formulate the dual

The various conditions were all given above, so concentrating on our lagrangian we have:

$$L(x,\lambda) = \frac{1}{2}x^T Gx + c^T x + \lambda^T(Ax-b)$$

Our dual equation will be:

$$g(\lambda) = \inf_x \left[\frac{1}{2}x^T Gx + c^T x + \lambda^T(Ax-b)\right]$$

and for our dual problem instead of minimising over x we now maximise over λ, in other words we seek $\max_\lambda g(\lambda)$.

Apologies I for some reason couldn't get latex to play ball with aligning the following cleanly.

If we substitute for x, we have:

$$g(\lambda) = \frac{1}{2}\left((G^{-1}(A^T\lambda + c))^T G(G^{-1}(A^T\lambda + c))\right) - c^T(G^{-1}(A^T\lambda+c)) - \lambda^T(A(G^{-1}(A^T\lambda+c))) - b^T\lambda$$

Therefore,

$$g(\lambda) = \frac{1}{2}((A^T\lambda+c)^T G^{-1}(A^T\lambda+c)) - c^T(G^{-1}(A^T\lambda+c)) - \lambda^T AG^{-1}(A^T\lambda+c) - b^T\lambda$$

$$g(\lambda) = \frac{1}{2}(\lambda^T AG^{-1}(A^T\lambda + c) + c^T G^{-1}(A^T\lambda + c)) - c^T(G^{-1}(A^T\lambda + c)) - \lambda^T AG^{-1}(A^T\lambda + c) - b^T\lambda$$

$$g(\lambda) = -\frac{1}{2}(\lambda^T AG^{-1}(A^T\lambda + c) + c^T G^{-1}(A^T\lambda + c)) - b^T\lambda$$

$$g(\lambda) = -\frac{1}{2}(A^T\lambda + c)^T G^{-1}(A^T\lambda + c) - b^T\lambda$$

So the dual problem is

$$\max_{\lambda} g(\lambda) = -\frac{1}{2}(A^T\lambda + c)^T G^{-1}(A^T\lambda + c) - b^T\lambda \text{ , subject to } \lambda \geq 0$$

With the usual interpretation for vector notation. Geometrically we have started with a convex minimisation problem and moved to a concave maximisation and where we have a solution we wind up approaching a saddle point in the space.

2 Exercise 2

$$\min_{x,y} f(x,y) = (x-2y)^2 + (x-2)^2 \text{ , subj to } x - y = 4$$

2.1 Formulate KKT system

- Lagrangian $L(x,y,\lambda) = (x-2y)^2 + (x-2)^2 + \lambda(x-y-4)$
- $x - y - 4 = 0$
- $\nabla_x L = 0 \implies 2(x-2y) + 2(x-2) + \lambda = 0$
- $\nabla_y L = 0 \implies -4(x-2y) - \lambda = 0$

2.2 Solve it

The approach in this case is to grab the three equations and solve as a linear system for λ, x and y.

- $x - y - 4 = 0$
- $4x - 4y - 4 + \lambda = 0$
- $-4x + 8y - \lambda = 0$

Here we can pretty much 'read off' the solution $y = 1$ and $x = 5$. We can also see that this means the λ constraint is kicking in as $\lambda = 0$ means that we would have the unconstrained minimum $x = 2$, $y = 1$, thus the equality constraint in this case is having an impact.