Norwegian University of Science and Technology
Department of Mathematical Sciences

# TMA4180 Optimisation I Spring 2017

**Solutions to exercise set 8**

**1** Constructively, a logarithmic barrier approach may be written as

$$\min_{x,y,s} x + y - \mu \log s \qquad \text{subject to} \qquad 1 - x^2 - y^2 - s = 0,$$

where $s$ ($\geq 0$) is the slack variable, and $\mu > 0$ is the barrier parameter which we intend to drive to 0. Introducing a Lagrange multipler $\lambda$, the KKT conditions for this problem are

$$1 + 2x\lambda = 0, \qquad 1 + 2y\lambda = 0, \qquad -\frac{\mu}{s} + \lambda = 0, \qquad \text{and} \qquad 1 - x^2 - y^2 - s = 0.$$

This gives first that

$$\lambda = \frac{\mu}{s} \qquad \text{and} \qquad x = y = -\frac{s}{2\mu},$$

and inserted into the constraint equation, we find that

$$1 - \frac{s^2}{2\mu^2} - s = 0.$$

The relevant solution of this quadratic equation is $s = \mu\big(\sqrt{\mu^2+2} - \mu\big)$, and we end up with

$$x = y = -\frac{1}{2}\big(\sqrt{\mu^2+2} - \mu\big) \qquad \text{and} \qquad \lambda = \big(\sqrt{\mu^2+2} - \mu\big)^{-1}.$$

Since the Hessian of the Lagrangian to this problem is positive definite (check it!), the found KKT point is the unique global minimiser of the logarithmic barrier formulation. Notably, as $\mu \to 0^+$, we recover the exact solution $x^* = y^* = -1/\sqrt{2}$, with $\lambda^* = 1/\sqrt{2}$, of the original problem.

**2** **a)** In matrix form—since we have $m$ constraints—the augmented Lagrangian equals

$$\mathcal{L}_A(x, \lambda; \mu) = \frac{1}{2}\|x\|^2 - \lambda^\top (Ax - b) + \frac{\mu}{2}\|Ax - b\|^2,$$

where $\lambda \in \mathbb{R}^m$ is the Lagrange multipler, and $\mu > 0$ is the penalty parameter. Trying to minimise $\mathcal{L}_A$ with respect to $x$, we first note that $\mathcal{L}_A(\cdot, \lambda; \mu)$ is both smooth and coercive, where the latter property follows from the dominating $\frac{1}{2}\|x\|^2$ term (how?). Hence, a global minimiser exists. This point is also a stationary point, and computing

$$0 = \nabla_x \mathcal{L}_A(x, \lambda; \mu) = x - A^\top \lambda + \mu A^\top (Ax - b)$$

shows that the unique global minimiser, for all $\lambda$ and all $\mu$, is

$$\begin{aligned} x_{\lambda,\mu} &= \Big(\frac{1}{\mu}\mathrm{Id} + A^\top A\Big)^{-1} A^\top \Big(\frac{1}{\mu}\lambda + b\Big) \\ &= A^\top \Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big)^{-1} \Big(\frac{1}{\mu}\lambda + b\Big). \end{aligned}$$

(Last transition was shown in exercise 3 b) in exercise set 7.)

**b)** Since the exact solution and optimal Lagrange multiplier of the original optimisation problem equal

$$x^* = (A^\top A)^{-1} A^\top b = A^\top (AA^\top)^{-1} b = A^\top \lambda^* \qquad \text{and} \qquad \lambda^* = (AA^\top)^{-1} b,$$

we demand that $x_{\lambda,\mu} = x^*$ and see what happens: left-multipling both sides by $(AA^\top)^{-1}A$ gives

$$\Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big)^{-1} \Big(\frac{1}{\mu}\lambda + b\Big) = \lambda^*,$$

or,

$$\frac{1}{\mu}\lambda + b = \Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big)\lambda^* = \frac{1}{\mu}\lambda^* + b.$$

In conclusion, the minimiser of the augmented Lagrangian equals that of the exact solution if and only if $\lambda = \lambda^*$, with no restrictions on $\mu > 0$.

**c)** Before we should make any use of the iterative algorithm, it is vital to establish its *consistency* with the original optimisation problem which it intends to solve. By this we mean that the algorithm should solve the original problem in the limit: if $x^k \to x$ and $\lambda^k \to \lambda$, then $x = x^*$ and $\lambda = \lambda^*$.
As $x^{k+1}$ is the minimum of $\mathcal{L}_A(\cdot, \lambda^k; \mu)$, we know that $x^{k+1} = x_{\lambda^k,\mu}$, and so inserting this into the iteration for the Lagrange multiplier yields that

$$\lambda^{k+1} = \lambda^k - \mu\left(Ax^{k+1} - b\right) = M(\lambda^k + \mu b),$$

where

$$\begin{aligned} M &= \mathrm{Id} - AA^\top \Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big)^{-1} \\ &= \left[\Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big) - AA^\top\right] \Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big)^{-1} \\ &= \frac{1}{\mu}\Big(\frac{1}{\mu}\mathrm{Id} + AA^\top\Big)^{-1} \\ &= \big(\mathrm{Id} + \mu AA^\top\big)^{-1}. \end{aligned}$$

Suppose now that $x^k \to x$ and $\lambda^k \to \lambda$. Then from the Lagrange multiplier iteration we get

$$\lambda = M(\lambda + \mu b),$$

which implies that

$$\big(\mathrm{Id} + \mu AA^\top\big)\lambda = \lambda + \mu b.$$

In other words,

$$\lambda = (AA^\top)^{-1}b = \lambda^*,$$

and therefore $x = x^*$ as well, because $x^*$ is the minimiser of $\mathcal{L}_A(\cdot, \lambda^*; \mu)$ from question b). Thus the scheme is consistent.

Recall now from numerical linear algebra that the consistent Lagrange multiplier iteration will converge for all initial values if and only if $\rho(M) < 1$, where $\rho(M)$ denotes the spectral radius of $M$, that is, the largest eigenvalue of $M$ in absolute value. Since $A$ has full rank, matrix $AA^\top$ is positive definite, with strictly positive eigenvalues. As such, the eigenvalues of

$$M^{-1} = \mathrm{Id} + \mu AA^\top$$

are always strictly greater than 1, which means that $\rho(M) < 1$, as desired.