SF2822 Applied nonlinear optimization, final exam
Friday June 3 2016 8.00–13.00
Brief solutions

1. We have

$$f(x)=\frac{1}{2}(x_1+1)^2+\frac{1}{2}(x_2+2)^2, \qquad g(x)=-2(x_1+x_2-1)^2-(x_1-x_2)^2+10,$$

$$\nabla f(x)=\begin{pmatrix} x_1+1 \\ x_2+2 \end{pmatrix}, \qquad \nabla g(x)=\begin{pmatrix} -6x_1-2x_2+4 \\ -2x_1-6x_2+4 \end{pmatrix},$$

$$\nabla^2 f(x)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \nabla^2 g(x)=\begin{pmatrix} -6 & -2 \\ -2 & -6 \end{pmatrix}.$$

(a) Insertion of numerical values in the expressions above gives the first QP-problem according to

$$\begin{array}{ll} \min & \frac{1}{2}p_1^2+\frac{1}{2}p_2^2+p_1+2p_2 \\ \text{subject to} & 4p_1+4p_2=-8. \end{array}$$

This is a convex QP-problem with a globally optimal solution given by

$$\begin{aligned} p_1-4\lambda &= -1, \\ p_2-4\lambda &= -2, \\ 4p_1+4p_2 &= -8. \end{aligned}$$

The solution is given by $p_1=-1/2$, $p_2=-3/2$ and $\lambda=1/8$, which agrees with the printout from the SQP-solver.

(b) We can see that $\nabla^2 f(x)$ and $-\nabla^2 g(x)$ are positive definite, independently of x. Moreover λ is nonnegative in all iterations. This implies that the solution to each QP subproblem is optimal also for the case when the equality constraint is relaxed to a greater than or equal constraint in (NLP). But, this is a convex problem since $\nabla^2 f(x)$ and $-\nabla^2 g(x)$ are positive definite, independently of x. Therefore, since the iterates converge to a point that satisfies the first-order necessary optimality conditions for the relaxed problem, and the relaxed problem is convex, it is a global minimizer of the relaxed problem and the original problem.

(c) We would expect quadratic rate of convergence, which is not seen. Therefore, we find it likely that something is not quite right. The first QP subproblem was correctly solved, and in this subproblem $\nabla^2 g(x)$ was of no importance, since $\lambda=0$. We therefore suspect that the evaluation of $\nabla^2 g(x)$ is not correct.

Remark: The iterates in the printout are generated with $\nabla^2 g(x)_{22}=-8$, which is not correct. If $\nabla^2 g(x)_{22}$ is correctly set to -6, the printout would be as follows.

It	x_1	x_2	λ	$\Vert\nabla f(x)-\nabla g(x)\lambda\Vert$	$\Vert g(x)\Vert$
0	0	0	0	2.2361	8.0000
1	-0.5000	-1.5000	0.1250	1.4577	9.0000
2	-0.2813	-1.0134	0.1004	0.0997	1.0668
3	-0.2382	-0.9444	0.1039	0.0022	0.0258
4	-0.2371	-0.9426	0.1044	$7.7122\cdot 10^{-6}$	$1.6646\cdot 10^{-5}$
5	-0.2371	-0.9426	0.1044	$8.3920\cdot 10^{-12}$	$6.9367\cdot 10^{-12}$
6	-0.2371	-0.9426	0.1044	$2.4825\cdot 10^{-16}$	0.0000

Now, quadratic rate of convergence is observed.

2. We may make use of the fact that the problem has only simple bounds.

For convenience, we write the constrainst in the form we normally use, $Ax \geq b$, with

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

Initially, $W^{(0)} = \{1, 2, 3\}$, so that three linearly active constraints are active, giving $p^{(0)} = 0$, so that $x^{(1)} = x^{(0)} = 0$. Therefore,

$$\begin{pmatrix} \lambda_1^{(1)} \\ \lambda_2^{(1)} \\ \lambda_3^{(1)} \end{pmatrix} = Hx^{(1)} + c = c = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix}.$$

Since $\lambda_1^{(1)} < 0$, constraint 1 is deleted from the working set. The search direction is given by

$$h_{11} p_1^{(2)} = -\lambda_1^{(2)}, \quad \text{i.e.} \quad 2p_1^{(2)} = 3,$$

so that $p^{(2)} = (3/2 \ 0 \ 0)^T$. We obtain $\alpha_{\max} = 2/3$, obtained for constraint 4, so that $\alpha^{(2)} = 2/3$, giving $x^{(3)} = (1 \ 0 \ 0)^T$ with $W^{(3)} = \{2, 3, 4\}$. Again, three linearly indepent active constraints are in the working set, giving $p^{(3)} = 0$, so that $x^{(4)} = x^{(3)}$. The multipliers are given by

$$\begin{pmatrix} -\lambda_4^{(4)} \\ \lambda_2^{(4)} \\ \lambda_3^{(4)} \end{pmatrix} = Hx^{(4)} + c = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$$

Since $\lambda_3^{(4)} < 0$, constraint 3 is deleted from the working set. The search direction is given by

$$h_{33} p_3^{(4)} = -\lambda_3^{(4)}, \quad \text{i.e.} \quad 3p_3^{(4)} = 1,$$

so that $p^{(4)} = (0 \ 0 \ 1/3)^T$. We obtain $\alpha_{\max} = 3$, so that $\alpha^{(4)} = 1$, giving $x^{(5)} = (1 \ 0 \ 1/3)^T$. The multipliers are given by

$$\begin{pmatrix} -\lambda_4^{(5)} \\ \lambda_2^{(5)} \\ 0 \end{pmatrix} = Hx^{(5)} + c = \begin{pmatrix} -\frac{5}{3} \\ \frac{4}{3} \\ 0 \end{pmatrix},$$

so that $\lambda_2^{(5)} = 4/3$ and $\lambda_4^{(5)} = \frac{5}{3}$. Since $\lambda^{(5)} \geq 0$, $x^{(5)}$ is optimal.

3. (a) The problem (QP) is a convex quadratic program.

The primal part of the trajectory is obtained as minimizer to the barrier-transformed problem

$$(P_\mu) \qquad \min \ \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 - \mu \ln(x_1 + x_2 + 1)$$

under the implicit condition that $x_1 + x_2 + 1 > 0$. The first-order optimality conditions of (P_μ) gives

$$\begin{aligned} x_1(\mu) - \frac{\mu}{x_1(\mu) + x_2(\mu) + 1} &= 0, \\ x_2(\mu) - \frac{\mu}{x_1(\mu) + x_2(\mu) + 1} &= 0. \end{aligned}$$

Subtraction of the second equation from the first gives $x_1(\mu) = x_2(\mu)$. Hence, we may let $x_1(\mu) = x_2(\mu) = t$. The equation then becomes

$$t - \frac{\mu}{2t+1} = 0,$$

so that

$$2t^2 + t - \mu = 0, \quad \text{i.e.,} \quad t^2 + \frac{t}{2} - \frac{\mu}{2} = 0, \quad \text{i.e.,}$$

Therefore

$$t = -\frac{1}{4} \pm \sqrt{\frac{1}{16} + \frac{\mu}{2}}.$$

The implicit constraint $2t + 1 > 0$ implies that the plus sign must be chosen, so that

$$x_1(\mu) = x_2(\mu) = -\frac{1}{4} + \sqrt{\frac{1}{16} + \frac{\mu}{2}}.$$

The dual part of the trajectory, i.e. $\lambda(\mu)$, is given by $\lambda_i(\mu) = \mu / g_i(x(\mu))$, $i = 1, \dots, m$. Here we only have one constraint, so

$$\lambda(\mu) = \frac{\mu}{2\left(-\frac{1}{4} + \sqrt{\frac{1}{16} + \frac{\mu}{2}}\right) + 1} = -\frac{1}{4} + \sqrt{\frac{1}{16} + \frac{\mu}{2}}.$$

(b) As $\mu \to 0$ it follows that $x(\mu) \to (0 \ 0)^T$ and $\lambda(\mu) \to 0$. Let $x^* = (0 \ 0)^T$ and $\lambda^* = 0$. Then x^* and λ^* satisfy the first-order optimality conditions of (QP). Since (QP) is a convex problem, this is sufficient for global optimality of (QP).

(c) We have

$$\lambda(\mu) - \lambda^* = -\frac{1}{4} + \sqrt{\frac{1}{16} + \frac{\mu}{2}} = \mu + o(\mu).$$

This is what we would expect, as $\lambda(\mu)$ for an inactive constraint is expected to be proportional to μ for small values of μ.

4. (See the course material.)

5. (a) We may write the Lagrangian function associated with (NLP_u) on the form

$$\mathcal{L}(x, u, \lambda, \eta) = f(x, u) - \lambda^T g(x, u) - \eta(u - \widetilde{u}),$$

where λ is the Lagrange multiplier vector associated with the constraint $g(x, u) \geq 0$ and η is the scalar Lagrange multiplier associated with the constraint $u - \widetilde{u} = 0$. The first-order optimality conditions then take the form

$$\nabla f(x, u) = \sum_{j=1}^{m} \lambda_i \nabla g_i(x, u), \tag{1a}$$

$$g_i(x, u) \geq 0, \quad i = 1, \ldots, m, \tag{1b}$$

$$\lambda_i \geq 0, \quad i = 1, \ldots, m, \tag{1c}$$

$$\lambda_i g_i(x, u) = 0, \quad i = 1, \ldots, m, \tag{1d}$$

$$\frac{\partial f(x, u)}{\partial u} = \sum_{j=1}^{m} \lambda_i \frac{\partial g_i(x, u)}{\partial u} + \eta, \tag{1e}$$

$$u = \widetilde{u}, \tag{1f}$$

where ∇ means taking derivatives with respect to x.

Then, (1a)–(1d) are the optimality conditions of (NLP), thus satisfied for $x = \widetilde{x}$, $\lambda = \tilde{\lambda}$ and $u = \widetilde{u}$. In addition, (1f) holds for $u = \widetilde{u}$, so that (1e) is satisfied for $\tilde{\eta}$ given by

$$\tilde{\eta} = \frac{\partial f(\widetilde{x}, \widetilde{u})}{\partial u} - \sum_{j=1}^{m} \tilde{\lambda}_i \frac{\partial g_i(\widetilde{x}, \widetilde{u})}{\partial u}. \tag{2}$$

(b) From sensititity analysis, we expect that the Lagrange multiplier associated with the constraint $u = \widetilde{u}$ predicts the change in optimal objective function value, so that

$$z(u) \approx z(\widetilde{u}) + \tilde{\eta}(u - \widetilde{u}) = z(\widetilde{u}) + \left(\frac{\partial f(\widetilde{x}, \widetilde{u})}{\partial u} - \sum_{j=1}^{m} \tilde{\lambda}_i \frac{\partial g_i(\widetilde{x}, \widetilde{u})}{\partial u} \right) (u - \widetilde{u}),$$

where the expression for $\tilde{\eta}$ given by (2) has been used. (This could be verified by the implicit function theorem with the assumptions given here.) Consequently, AF is not correct. He has missed that the dependence on u in the constraint functions is also of importance.