SF2822 Applied nonlinear optimization, final exam
Thursday May 30 2024 8.00–13.00
Brief solutions

1. (a) GAMS has terminated successfully with model status ``Locally Optimal'', meaning that a point $x^*$ and Lagrange multiplier vector $\lambda^*$ that together satisfy the first-order necessary optimality conditions for $(NLP)$ have been computed. From ``LEVEL'' of ``VAR x'', we obtain the solution

$$x^* \approx \begin{pmatrix} -1.000 & 0.000 & -0.707 \end{pmatrix}^T.$$

Analogously, the Lagrange multipliers of the constraints are given by ``MARGIN'' of ``EQU cons1'', ``EQU cons2'', and ``VAR x'' associated with ``j1'' and ``j2'', as

$$\lambda^* \approx \begin{pmatrix} 0.000 & 45.752 & 24.488 & 285.643 \end{pmatrix}^T.$$

(b) We have $f(x) = f_1(x_1 + 2x_2 + x_3 + 5) + f_2(2x_1 + x_3 - 4)$ for $f_1(y) = y^4$ and $f_2(y) = y^2$. Then, $f_1''(y) = 12y^2 \geq 0$ and $f_2''(y) = 2 \geq 0$, so that f_1 and f_2 are convex functions on $I\!R$. As linear tranformations preserve convexity, we obtain f as a sum of two convex functions, hence convex. As $g_1(x^*) > 0$, constraint one is not active at x^*, so we may consider the problem

$$(NLP') \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & -x_1^2 - 2x_3^2 + 2 = 0, \\ & x_1 + 1 \geq 0, \\ & x_2 \geq 0. \end{array}$$

Problem (NLP) is not convex, due to a nonlinear equality constraint. However, as $\lambda_2^* \geq 0$, it follows that x^* together with λ^* satisfy the first-order optimality conditions for the problem (NLP''), where the constraint $-x_1^2 - 2x_3^2 + 2 = 0$ is replaced by $-x_1^2 - 2x_3^2 + 2 \geq 0$, i.e.,

$$(NLP'') \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & -x_1^2 - 2x_3^2 + 2 \geq 0, \\ & x_1 + 1 \geq 0, \\ & x_2 \geq 0. \end{array}$$

Now, $g_2(x) = -x_1^2 - 2x_3^2 + 2$ is a concave function. Therefore, the feasible region of (NLP'') is convex, so that x^* is a global minimizer to (NLP''). As (NLP'') is a relaxation of (NLP), x^* is a global minimizer to (NLP) as well.

(A less insightful answer is that global optimality cannot be concluded directly from the first-order necessary optimality conditions as (NLP) is not a convex problem.)

(c) The expected change in the objective function is given by the Lagrange multiplier, up to first order, hence $162.5591 - 45.752t$ for the second constraint and $162.5591 - 24.488t$ for the third constraint. Therefore, the second constraint is to be preferred.

2. (a) Problem (QP) is a convex quadratic program with objective function $\frac{1}{2}x^T Hx + c^T x$ for $H = I$ and $c = 0$, and constraint $Ax \geq b$ for $A = (1\ 1)$ and $b = 2$. The primal-dual nonlinear equations take the form

$$Hx + c - A^T\lambda = 0$$
$$(Ax - b)\lambda - \mu = 0.$$

Insertion of numerical values gives

$$\begin{aligned} 3x_1 - \lambda &= 0, \\ x_2 - \lambda &= 0, \\ (x_1 + x_2 - \frac{8}{3})\lambda - \mu &= 0. \end{aligned}$$

We may express x_1 and x_2 in λ from the first two equations as

$$x_1 = \frac{\lambda}{3}, \quad x_2 = \lambda,$$

which inserted in the third equation gives

$$(\frac{\lambda}{3} + \lambda - \frac{8}{3})\lambda - \mu = 0.$$

This is equivalent to

$$\lambda^2 - 2\lambda - \frac{3}{4}\mu = 0.$$

Therefore,

$$\lambda(\mu) = 1 + \sqrt{1 + \frac{3}{4}\mu}.$$

The plus sign is chosen as $\lambda(\mu) > 0$. Then,

$$x_1(\mu) = \frac{\lambda(\mu)}{3} = \frac{1}{3}\left(1 + \sqrt{1 + \frac{3}{4}\mu}\right), \quad x_2(\mu) = \lambda(\mu) = 1 + \sqrt{1 + \frac{3}{4}\mu}.$$

(As a check, we may verify $\lim_{\mu\to 0+} x(\mu) = (2/3 \;\; 2)^T = x^*$ and $\lim_{\mu\to 0+} \lambda(\mu) = 2 = \lambda^*$.)

(b) The Newton step Δx, $\Delta\lambda$ is given by linearization of the primal-dual nonlinear equations as

$$\begin{pmatrix} H & -A^T \\ \lambda A & Ax - b \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} Hx + c - A^T\lambda \\ (Ax - b)\lambda - \mu \end{pmatrix},$$

where the right-hand side is evaluated at the particular iterate x, λ. Insertion of numerical values gives

$$\begin{pmatrix} 3 & 0 & -1 \\ 0 & 1 & -1 \\ 2 & 2 & \frac{4}{3} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta\lambda \end{pmatrix} = \begin{pmatrix} -10 \\ 2 \\ -\frac{5}{3} \end{pmatrix}.$$

3. If the problem is put on the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \ge 0, \quad x \in I\!R^2, \end{array}$$

we obtain

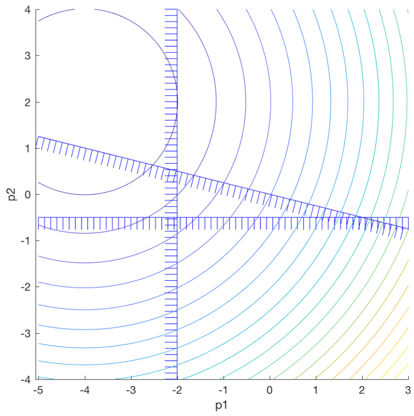
$$\nabla f(x)^T = \begin{pmatrix} x_1 + x_2 + \frac{3}{2} & x_1 + x_2 - \frac{9}{2} \end{pmatrix}, \quad \nabla g(x)^T = \begin{pmatrix} x_2 & x_1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \begin{pmatrix} 1 & 1 - \lambda_1 \\ 1 - \lambda_1 & 1 \end{pmatrix}.$$

With $x^{(0)} = (2\ \frac{1}{2})^T$ and $\lambda^{(0)} = (1\ 0\ 0)^T$, the first QP-problem becomes

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\begin{pmatrix} p_1 & p_2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} 4 & -2 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\ \text{subject to} & \begin{pmatrix} \frac{1}{2} & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ -2 \\ -\frac{1}{2} \end{pmatrix}. \end{array}$$

The optimal solution of the QP-problem is given by the feasible point which is closest, in 2-norm, to $(-4\ \ 2)^T$. This may for example be solved graphically:



The solution is $p^{(0)} = (-2\ \ 2)^T$ with constraint 2 active. The Lagrange multiplier $\lambda_2^{(1)}$ of the active constraint is given by

$$\begin{pmatrix} -2 \\ 2 \end{pmatrix} + \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \lambda_2^{(1)},$$

i.e., $\lambda_2^{(1)} = 2$. Thus, we have $\lambda^{(1)} = (0\ \ 2\ \ 0)^T$, and $x^{(1)}$ is given by $x^{(1)} = x^{(0)} + p^{(0)} = (0\ 5/2)^T$.

4. (See the course material.)

5. (a) We may let $M = \sum_{i=1}^n U_i w_i - V$. Then, $M - tI \succeq 0$ ensures $t \leq \eta_{\min}(M)$ and $sI - M \succeq 0$ ensures $s \geq \eta_{\max}(M)$. We may then formulate the problem as

$$(P) \quad \begin{array}{ll} \underset{s,t,w}{\text{minimize}} & s - t \\ \text{subject to} & sI - \sum_{i=1}^n U_i w_i + V \succeq 0, \\ & \sum_{i=1}^n U_i w_i - V - tI \succeq 0. \end{array}$$

(b) We may derive the dual by Lagrangian relaxation. Let Y and Z be positive semidefinite symmetric $m \times m$ matrices. Then,

$$\begin{aligned} L(s,t,w,Y,Z) &= s - t - \text{trace}(Y(sI - \sum_{i=1}^n U_i w_i + V)) - \text{trace}(Z(\sum_{i=1}^n U_i w_i - V - tI)) \\ &= \text{trace}(VZ) - \text{trace}(VY) + s(1 - \text{trace}(Y)) - t(1 - \text{trace}(Z)) \end{aligned}$$

$$+\sum_{i=1}^{n} w_i(\text{trace}(U_iY) - \text{trace}(U_iV)).$$

We obtain

$$\min_{s,t,w} L(s,t,w,Y,Z) = \begin{cases} \text{trace}(VZ) - \text{trace}(VY) & \text{if} \quad 1 - \text{trace}(Y) = 0, 1 - \text{trace}(Z) = 0, \\ & \qquad \text{trace}(U_iY) - \text{trace}(U_iV) = 0, i = 1, \ldots, n, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is therefore obtained as

$$(D) \qquad \begin{array}{ll} \underset{Y,Z}{\text{maximize}} & \text{trace}(VZ) - \text{trace}(VY) \\ \text{subject to} & \text{trace}(Y) = 1, \\ & \text{trace}(Z) = 1, \\ & \text{trace}(U_iY) = \text{trace}(U_iV), \quad i = 1, \ldots, n, \\ & Y = Y^T \succeq 0, \\ & Z = Z^T \succeq 0. \end{array}$$

(The dual could equivalently be derived by reformulating the primal problem to the form given in the hint and use the corresponding dual problem.)