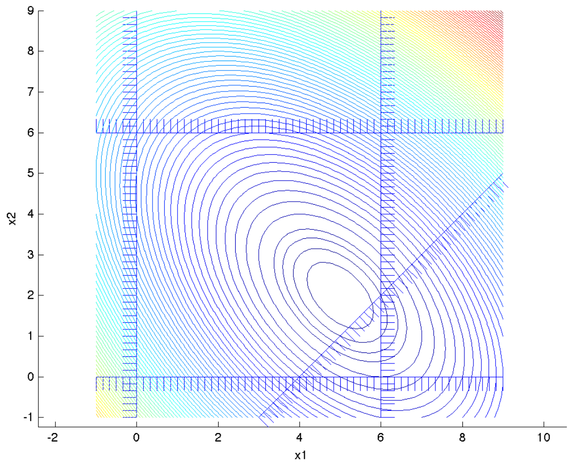
SF2822 Applied nonlinear optimization, final exam
Thursday June 2 2022 8.00–13.00
Brief solutions

---

1. (a) The iterations are illustrated in the figure below:



In the first iteration the search direction points at $(6\ 1.5)^T$, but the step is limited by the constraint $-x_1 + x_2 \geq -4$, which is added so that the new point is $(6\ 2)^T$. A zero step is taken, and the multiplier for the constraint $x_2 = 6$ is negative. Thus, this constraint is deleted. The new step points at $(11/2\ 3/2)$, which is feasible. A unit step is taken, and the multiplier for $-x_1 + x_2 = -4$ is negative, $-1/2$. This constraint is deleted. The new step points at $(5\ 2)^T$, which is feasible. No constraints are active, so this point is optimal.

(b) In the first iteration the search direction points at $(5\ 2)^T$, but the step is limited by the constraint $-x_1 \geq -4$, which is added, and the new point is $(4\ 10/3)$. The new step points at $(4\ 5/2)$, which is feasible. The multiplier of the constraint is positive, $3/2$, so that an optimal solution has been found.

2. (See the course material.)

3. (a) The problem is convex as it can equivalently be stated as the unconstrained convex quadratic problem $\min -x_1(2 - x_1) = \min x_1^2 - 2x_1$. The primal part of the trajectory is obtained as minimizer to the quadratically-penalized problem

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

The first-order optimality conditions of (P_μ) gives

$$-x_2(\mu) - \frac{1}{\mu}(x_1(\mu) + x_2(\mu) - 2) = 0,$$
$$-x_1(\mu) - \frac{1}{\mu}(x_1(\mu) + x_2(\mu) - 2) = 0.$$

These equations are symmetric in $x_1(\mu)$ and $x_2(\mu)$. Hence, $x_1(\mu) = x_2(\mu)$. This means that $-x_1(\mu) - \frac{1}{\mu}(2x_1(\mu) - 2) = 0$, from which it follows that

$$x_1(\mu) = x_2(\mu) = \frac{2}{-\mu + 2}.$$

Since (P_μ) is a convex problem, this is a global minimizer.
The dual part of the trajectory, i.e. $\lambda(\mu)$, is normally given by $\lambda_i(\mu) = -g_i(x(\mu))/\mu$, $i = 1, \ldots, m$. Here we only have one constraint, so

$$\lambda(\mu) = -\frac{x_1(\mu) + x_2(\mu) - 2}{\mu} = -\frac{4/(2-\mu) - 2}{\mu} = -\frac{4}{(2-\mu)\mu} + \frac{2(2-\mu)}{\mu(2-\mu)} = \frac{-2\mu}{\mu(2-\mu)} = \frac{-2}{2-\mu}.$$

(b) As $\mu \to 0$ it follows that $x(\mu) \to (1\ 1)^T$ and $\lambda(\mu) \to -1$. Let $x^* = (1\ 1)^T$ and $\lambda^* = 1$. 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

$$x_1(\mu) - x_1^* = x_2(\mu) - x_2^* = \frac{2}{2-\mu} - 1 = \frac{\mu}{2-\mu} = \frac{1}{2}\mu + o(\mu).$$

This is as expected. We would expect $\|x(\mu) - x^*\|_2$ to be of the order μ near an optimal solution where regularity holds.

4. We have

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

(a) The point $x = (2,3)^\top$ is not feasible and thus JR is wrong.

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

$$\begin{array}{ll} \min & \frac{3}{2}p_1^2 + \frac{3}{2}p_2^2 - 2p_1 - 2p_2 \\ \text{subject to} & -p_2 \ge -\frac{1}{2}. \end{array}$$

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

$$\begin{aligned} 3p_1 - 2 &= 0, \\ 3p_2 - 2 + \lambda &= 0, \\ -p_2 &= -\frac{1}{2}, \end{aligned}$$

where the constraint is active since otherwise $\lambda = 0$ would lead to $p_2 = \frac{2}{3}$, which is infeasible. The solution is given by $p_1 = 2/3$, $p_2 = 1/2$ and $\lambda = 1/2$. Hence,

$$x^{(1)} = x^{(0)} + p = \begin{pmatrix} 2/3 \\ 3/2 \end{pmatrix}, \quad \lambda^{(1)} = \lambda = 1/2.$$

5. (a) Consider $v_1, \ldots, v_k \in \mathbb{R}^n$ be the orthonormal eigenvectors of $\lambda_1, \ldots, \lambda_k$ respectively. Then, the matrix $Y = \begin{pmatrix} v_1 & \ldots & v_k \end{pmatrix} \in \mathbb{R}^{n\times k}$ is feasible for μ_1 since

$Y^\top Y = I_k$ holds as the eigenvectors are orthonormal. Moreover, it yields an objective value of

$$\begin{aligned}\mathrm{trace}(CYY^\top) &= \mathrm{trace}(Y^\top CY) = \mathrm{trace}(Y^\top \begin{pmatrix} \lambda_1 v_1 & \ldots & \lambda_k v_k \end{pmatrix}) \\ &= \mathrm{trace}(\quad \mathrm{Diag} \quad (\lambda_1, \ldots, \lambda_k)) = \lambda_1 + \ldots + \lambda_k.\end{aligned}$$

(b) We observe that the matrix $Y = \begin{pmatrix} v_1 & \ldots v_k \end{pmatrix} \in \mathbb{R}^{n\times k}$ is also feasible for μ_2 as it is:

- positive semidefinite,
- $\mathrm{trace}(YY^\top) = \mathrm{trace}(Y^\top Y) = \mathrm{trace}(I_k) = k$ and
- $I_n - YY^\top \succeq 0$ since:

Consider an arbitrary vector $w = \sum_{i=1}^n \eta_i v_i$, then

$$\begin{aligned}w^\top \left(I_n - YY^\top\right) w \qquad &= w^\top w - (Y^\top w)^\top Y^\top w \\ &= (\sum_{i=1}^n \eta_i v_i)^\top (\sum_{i=1}^n \eta_i v_i) - (\sum_{i=1}^k \eta_i e_i)^\top (\sum_{i=1}^k \eta_i e_i) \\ &= \sum_{i,j=1}^n \eta_i \eta_j v_i^\top v_j - \sum_{i=1}^k \eta_i^2 = \sum_{i=k+1}^n \eta_i^2 \geq 0.\end{aligned}$$

Lastly, we know from (a) that $\mathrm{trace}(CYY^\top) = \lambda_1 + \ldots + \lambda_k$.

(c) The dual program of

$$\max_X \{\mathrm{trace}(CX) : \mathrm{trace}(X) = k, X \succeq 0\}$$

is

$$\min_{y\in\mathbb{R}} \{ky : yI_n - C \succeq 0\}.$$