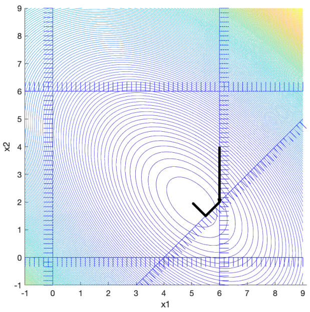
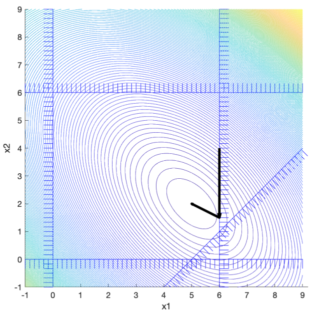
SF2822 Applied nonlinear optimization, final exam
Wednesday August 20 2025 8.00–13.00
Brief solutions

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



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

(b) The iterations are illustrated in the figure below:



In the first iteration the search direction points at $(6\ 3/2)^T$, which is feasible, so that a unit step is taken. The multiplier of the constraint $-x_1 = -6$ is negative, $-3/2$. This constraint is deleted from the working set. The new step points at $(5\ 2)^T$, which is feasible. A unit step is taken. No constraints are active, so this point is optimal.

2. (a) Problem (QP) is a convex quadratic program.
The primal part of the trajectory is obtained as minimizer to the barrier-transformed problem

$$(P_\mu) \qquad \min \quad \tfrac{1}{2}x_1^2 + x_2^2 - \mu \ln(x_1 + x_2 - 3)$$

under the implicit condition that $x_1 + x_2 - 3 > 0$. The first-order optimality conditions of (P_μ) gives

$$x_1(\mu) - \frac{\mu}{x_1(\mu) + x_2(\mu) - 3} = 0,$$
$$2x_2(\mu) - \frac{\mu}{x_1(\mu) + x_2(\mu) - 3} = 0.$$

Subtraction of the second equation from the first gives $x_1(\mu) = 2x_2(\mu)$. Hence, we may let $x_2(\mu) = t$, $x_1(\mu) = 2t$. The equation then becomes

$$2t - \frac{\mu}{3t-3} = 0,$$

so that

$$6t^2 - 6t - \mu = 0, \quad \text{i.e.,} \quad t^2 - t - \frac{\mu}{6} = 0.$$

Therefore

$$t = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{\mu}{6}}.$$

The implicit constraint $3t - 3 > 0$ implies that the plus sign must be chosen, so that

$$x_1(\mu) = 1 + 2\sqrt{\frac{1}{4} + \frac{\mu}{6}}, \quad x_2(\mu) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\mu}{6}}.$$

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

$$\lambda(\mu) = \frac{\mu}{3\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\mu}{6}}\right) - 3} = 1 + 2\sqrt{\frac{1}{4} + \frac{\mu}{6}} = 1 + \sqrt{1 + \frac{2\mu}{3}}.$$

(b) As $\mu \to 0$ it follows that $x(\mu) \to (2\ 1)^T$ and $\lambda(\mu) \to 2$. Let $x^* = (2\ 1)^T$ and $\lambda^* = 2$. 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

$$\lambda(\mu) - \lambda^* = -1 + \sqrt{1 + \frac{2\mu}{3}} = \frac{\mu}{3} + o(\mu).$$

This is what we would expect. We have a regular point at which strict complementarity holds and then $\lambda(\mu) - \lambda^*$ is expected to be proportional to μ for small values of μ.

3. (a) Since (NLP') is formed by perturbing the first constraint of (NLP) from $h(x) \geq 0$ to $h(x) \geq 1/4$, sensitivity analysis gives the estimate

$$f(\widetilde{x}) + \frac{1}{4}\tilde{\lambda}_1 = f(\widetilde{x}) + \frac{1}{2} = \frac{11}{2}.$$

(b) The QP-subproblem takes the general form

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}p^T \nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)})p + \nabla f(x^{(0)})^T p \\ \text{subject to} & A(x^{(0)})p \geq -g(x^{(0)}). \end{array}$$

We obtain

$$\begin{aligned} \nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)}) &= \nabla^2 f(x^{(0)}) - \lambda_1^{(0)} \nabla^2 h(x^{(0)}) \\ &= \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} - 2\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}, \end{aligned}$$

$$g(x^{(0)}) = \begin{pmatrix} h(x^{(0)}) - \frac{1}{4} \\ x^{(0)} \end{pmatrix} = \begin{pmatrix} -\frac{1}{4} \\ 2 \\ 2 \end{pmatrix},$$

$$A(x^{(0)}) = \begin{pmatrix} \nabla h(x^{(0)})^T \\ I \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Insertion of numerical values gives

$$\begin{array}{ll} \text{minimize} & 3p_1^2 + 2p_2^2 + 2p_1 \\ \text{subject to} & p_1 \geq \frac{1}{4}, \\ & p_1 \geq -2, \\ & p_2 \geq -2. \end{array}$$

This is a separable problem, so that minimization can be done with respect to p_1 and p_2 independently. We obtain $p_1 = 1/4$ and $p_2 = 0$ with Lagrange multipliers $\lambda_1 = 7/2$, $\lambda_2 = 0$ and $\lambda_3 = 0$. Consequently,

$$x^{(1)} = x^{(0)} + p = \begin{pmatrix} \frac{9}{4} \\ 2 \end{pmatrix}, \quad \lambda^{(1)} = \lambda = \begin{pmatrix} \frac{7}{2} \\ 0 \\ 0 \end{pmatrix}.$$

4. (See the course material.)

5. The second-order sufficient optimality conditions for (NLP_1) imply that
(i) $g(x^*) \geq 0$,
(ii) $\nabla f(x^*) = A(x^*)^T \lambda^*$ for some $\lambda^* \geq 0$,
(iii) $\lambda_i^* g_i(x^*) = 0$, $i = 1, \ldots, m$, and
(iv) $Z_+(x^*)^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) Z_+(x^*) \succ 0$,
where $A_+(x^*)$ contains the rows of $A(x^*)$ for which λ^* has positive components, and $Z_+(x^*)$ is a matrix whose columns form a basis for null$(A_+(x^*))$.

We may write (NLP_2) as

$$(NLP_2) \qquad \begin{array}{ll} \text{minimize} & \widetilde{f}(z, x) \\ \text{subject to} & \widetilde{g}(z, x) \geq 0, \end{array}$$

with

$$\widetilde{f}(z, x) = z, \quad \widetilde{g}(z, x) = \begin{pmatrix} z - f(x) \\ g(x) \end{pmatrix}.$$

Associated with (NLP_2), we may define the Lagrangian function

$$\tilde{\mathcal{L}}(z, x, \mu, \eta) = z - \mu(z - f(x)) - \eta^T g(x),$$

where μ is the Lagrange multiplier associated with $z - f(x) \geq 0$ and η are the Lagrange multipliers associated with $g(x) \geq 0$.

We now want to find z^*, μ^* and η^* so that the second-order sufficient optimality conditions (i)–(iv) hold, but associated with (NLP_2). This means that we want to find z^*, μ^* and η^* such that

(i') $\begin{pmatrix} z^* - f(x^*) \\ g(x^*) \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

(ii') $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\nabla f(x^*) & A(x^*)^T \end{pmatrix} \begin{pmatrix} \mu^* \\ \eta^* \end{pmatrix}$ for some $\mu^* \geq 0$ and $\eta^* \geq 0$,

(iii') $\mu^*(z^* - f(x^*)) = 0$, $\eta_i^* g_i(x^*) = 0$, $i = 1, \dots, m$, and

(iv') $\tilde{Z}_+(z^*, x^*)^T \nabla^2_{z,x} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \tilde{Z}_+(z^*, x^*) \succ 0$,

where $\tilde{Z}_+(z^*, x^*)$ is a matrix whose columns form a basis for $\text{null}(\tilde{A}_+(z^*, x^*))$, with $\tilde{A}_+(z^*, x^*)$ defined as the matrix comprising the rows of

$$\begin{pmatrix} 1 & -\nabla f(x^*)^T \\ 0 & A(x^*) \end{pmatrix}$$

for which the associated components of the multipliers μ^* and η^* of (ii') are positive.

We now verify these conditions. For (i') to hold, we must have $z^* \geq f(x^*)$, since $g(x^*) \geq 0$ holds by (i).

For (ii'), the first equation reads $1 = \mu^*$. Hence, $\mu^* = 1$ must hold. With $\mu^* = 1$, the second block of equations reads

$$0 = -\nabla f(x^*) + A(x^*)^T \eta^*,$$

which holds for $\eta^* = \lambda^*$ by (ii). Since $\mu^* = 1 > 0$ and $\lambda^* \geq 0$ by (ii), (ii') holds.

Since $\mu^* > 0$, (iii') holds if $z^* = f(x^*)$, since (iii) implies that $\eta_i^* g_i(x^*) = 0$, $i = 1, \dots, m$, if $\eta^* = \lambda^*$. In addition, since $z^* = f(x^*)$, (i') holds.

Finally, to verify (iv'), taking the derivatives gives

$$\nabla^2_{z,x} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) = \begin{pmatrix} \nabla^2_{zz} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) & \nabla^2_{zx} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \\ \nabla^2_{xz} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) & \nabla^2_{xx} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) \end{pmatrix}.$$

Since $\mu^* > 0$ and $\eta^* = \lambda^*$, we obtain

$$\tilde{A}_+(z^*, x^*) = \begin{pmatrix} 1 & -\nabla f(x^*)^T \\ 0 & A_+(x^*) \end{pmatrix} = \begin{pmatrix} 1 & -\lambda_+^{*\,T} A_+(x^*) \\ 0 & A_+(x^*) \end{pmatrix}.$$

Note that $\text{rank}(\tilde{A}_+(z^*, x^*)) = \text{rank}(A_+(x^*)) + 1$, since the first row of $\tilde{A}_+(z^*, x^*))$ is not linearly dependent on the other rows. Hence, $\text{null}(\tilde{A}_+(z^*, x^*))$ and $\text{null}(A_+(x^*))$ have the same dimension. Since

$$\begin{pmatrix} 1 & -\lambda_+^{*\,T} A_+(x^*) \\ 0 & A_+(x^*) \end{pmatrix} \begin{pmatrix} 0 \\ Z_+(x^*) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{we may let} \quad \tilde{Z}_+(z^*, x^*) = \begin{pmatrix} 0 \\ Z_+(x^*) \end{pmatrix}.$$

Then,

$$\begin{aligned} & \tilde{Z}_+(z^*, x^*)^T \nabla^2_{z,x} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \tilde{Z}_+(z^*, x^*) \\ &= \begin{pmatrix} 0 & Z_+(x^*)^T \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) \end{pmatrix} \begin{pmatrix} 0 \\ Z_+(x^*) \end{pmatrix} \\ &= Z_+(x^*)^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) Z_+(x^*) \succ 0, \end{aligned}$$

as required, where (iv) has been used in the last step. This means that the second-order sufficient optimality conditions hold for (NLP_2) with $z^* = f(x^*)$, $\mu^* = 1$ and $\eta^* = \lambda^*$.