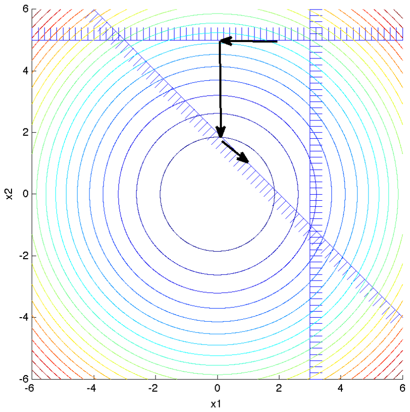
KTH Mathematics

SF2822 Applied nonlinear optimization, final exam
Tuesday August 23 2011 14.00–19.00
Brief solutions

---

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

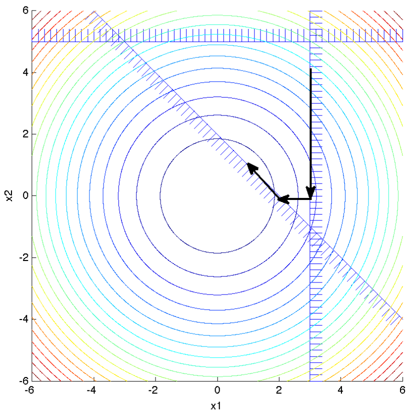


In the first iteration, the search direction points towards the minimizer with constraint 3 active, $(0\ 5)$, which is feasible. Hence, the next iterate is $(0\ 5)$. Constraint 3 has a negative multiplier, so it is deleted.

In the second iteration, the search direction points towards the unconstrained minimizer, $(0\ 0)$, but the step is limited at $(0\ 2)$ by constraint 1. Hence, constraint 1 is added.

In the third iteration, the search direction points towards the minimizer with constraint 1 active, $(1\ 1)$, which is feasible. Hence, the next iterate is $(1\ 1)$. Constraint 1 has a positive multiplier, so the optimal solution has been found.

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



In the first iteration, the search direction points towards the minimizer with constraint 2 active, $(3\ 0)$, which is feasible. Hence, the next iterate is $(3\ 0)$. Constraint 2 has a negative multiplier, so it is deleted.

In the second iteration, the search direction points towards the unconstrained minimizer, (0 0), but the step is limited at (2 0) by constraint 1. Hence, constraint 1 is added.
In the third iteration, the search direction points towards the minimizer with constraint 1 active, (1 1), which is feasible. Hence, the next iterate is (1 1). Constraint 1 has a positive multiplier, so the optimal solution has been found.

2. Since $g(x^{(0)}) > 0$, it is not necessary to introduce slack variables for the constraints. If slack variables are not introduced, the Newton step Δx, $\Delta\lambda$ is given by

$$\begin{pmatrix} \nabla^2_{xx}\mathcal{L}(x,\lambda) & -A(x)^T \\ \Lambda A(x) & G(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ G(x)\lambda - \mu e \end{pmatrix},$$

where $G(x) = \operatorname{diag}(g(x))$, $\Lambda = \operatorname{diag}(\lambda)$ and e is the vector of ones.
In our case we get

$$\begin{pmatrix} 1+2\lambda & 0 & 2x_1 \\ 0 & 1+\lambda & x_2 \\ -2\lambda x_1 & -\lambda x_2 & 2 - x_1^2 - \frac{1}{2}x_2^2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} x_1 - 3 + 2\lambda x_1 \\ x_2 - 2 + \lambda x_2 \\ (2 - x_1^2 - \frac{1}{2}x_2^2)\lambda - \mu \end{pmatrix}.$$

Then, for the first iteration we obtain

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

The next iterate is given by $x^{(1)} = x^{(0)} + \alpha^{(0)}\Delta x_1^{(0)}$, $\lambda^{(1)} = \lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)}$, where $\alpha^{(0)}$ is given by some approximate linesearch. The steplength $\alpha^{(0)}$ must be chosen such that $g(x^{(0)} + \alpha^{(0)}\Delta x^{(0)}) > 0$ and $\lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)} > 0$.

3. (a) We may write $A = (B\ N)$, where $B = I$ and $N = -e$, where e is the vector of ones. Then, a matrix Z whose columns form a basis for the nullspace of A is given by

$$Z = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} = \begin{pmatrix} e \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

(b) The step to the minimizer is given by Zp_Z, where

$$Z^T H Z p_Z = -Z^T(H\bar{x} + c).$$

Insertion of numerical values gives $6p_Z = 6$, i.e., $p_Z = 1$. Hence the optimal x is given by

$$x = \bar{x} + Zp_Z = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

The Lagrange multipliers are then given by $Hx + c = A^T\lambda$, i.e.,

$$\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix},$$

i.e., $\lambda = (1 \;\; 0 \;\; -1)^T$.

(c) Since $\lambda_2 = 0$, the first-order necessary optimality conditions will be fulfilled even if the second constraint is omitted. As we have a convex problem, the first-order necessary optimality conditions are sufficient for global optimality. Hence, the given point remains optimal to the problem where the second constraint is removed.

4. (See the course material.)

5. (a) The matrix D has three negative eigenvalues. Since the problem has two constraints, at most two constraints can be active at any feasble point. At least three constraints would have to be active for the reduced Hessian of the objective function with respect to the active constraints to be positive semidefinite. This is a necessary condition for a local minimizer. Hence, no local minimizer can exist.

(b) As outlined in Exercise 5a, since D is diagonal with $D_{ii} < 0$, $i = 1, 3, 5$, there must be at least three active constraints at any local minimizer. Hence, it must hold that $Ax^* = b$, and in addition at least one more constraint must be active, which is linearly independent from the rows of A. Since $Ae_1 = 0$ and $e_1^T De_1 < 0$, it follows that it is necessary to add a bound constraint of the form either $-x_1 \geq -1$ or $x_1 \geq 1$. If such a constraint is added, a matrix whose columns gives a basis for the resulting nullspace is given by A matrix whose columns form a basis for the nullspace of A is given by

$$Z = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},$$

so that

$$Z^T DZ = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix},$$

which is a positive definite matrix. We assume that the bound-constraint $x_1 \geq 1$ is added, and try to establish the existence of Lagrange multipliers. The requirement becomes $Dx^* + c = A^T\lambda + e_1\lambda_3$, where $\lambda = (\lambda_1\ \lambda_2)^T$. Thus,

$$\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix},$$

which has a solution $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = -1$. Since $\lambda_3 < 0$ we conclude that we should instead add the bound-constraint $-x_1 \geq -1$, which changes the sign of λ_3 in the resulting problem. Then, x^* satisfies the second-order sufficient conditions for a local minimizer to the resulting problem. Hence, x^* becomes a local minimizer if the constraint $-x_1 \geq -1$ is added to the problem.