NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF ENGINEERING CYBERNETICS

Contact for questions during exam:
Name: Professor Bjarne Foss
Tel: 92422004

English version

Exam in TTK4135

Optimization and Control

Optimalisering og regulering

Thursday June 10, 2010

Duration: 0900 - 1300

Combination of allowed help remedies:
D - No printed or hand-written notes.
Certified calculator with empty memory.

In the Appendix potentially useful information is included.
The grades will be available by July 1.

1 Nelder-Mead method (30%)

a What is meant by the term a derivative-free optimization algorithm?

b Assume that you are searching for the deepest point in a lake and measure the depth with an echo sounder placed in a boat. You choose the Nelder-Mead algorithm (N-M) for this purpose. The feasible set is a subset of $x \in \mathbb{R}^n$ given by the square $X = \left\{x \in \mathbb{R}^2, \quad 0 \le x_1 \le 1, \quad 0 \le x_2 \le 1\right\}$. N-M needs three points to proceed. The water depth is given by $f(x)$. Assume the following ordered points.

$$\begin{aligned} x^1 &= (0.7, 0.8)^T, \quad f(x^1) = 15 \\ x^2 &= (0.5, 0.8)^T, \quad f(x^2) = 20 \\ x^3 &= (0.7, 0.5)^T, \quad f(x^3) = 30 \end{aligned}$$

The reflection point x^{refl} for $x \in R^n$ is given by

$$x^{refl} \stackrel{def}{=} g(1)$$

where

$$\begin{aligned} g(t) &= \overline{x} + t(x^{n+1} - \overline{x}) \\ \overline{x} &= \frac{1}{n}\sum_{i=1}^{n} x^i \end{aligned}$$

Compute the reflection point for our problem and show it graphically together with the other points (x^1, x^2, x^3) and the feasible set.

The points x^1, x^2, x^3 have been ordered in a specific way. Explain this ordering.

c Show how N-M computes the next iteration point when the water depth at the reflection point is $f(x^{refl}) = 35$.

Show how N-M computes the next iteration point when the water depth at the reflection point is $f(x^{refl}) = 25$.

Show how N-M computes the next iteration point when the water depth at the reflection point is $f(x^{refl}) = 10$.

Please present your results in the three cases through three independent macro codes.

d In the last case where $f(x^{refl}) = 10$ we need to apply "inside" contraction meaning that the original triangle defined by x^1, x^2, x^3 is reduced in size. Assume further that $f(g(t = \frac{1}{2})) = 18$. Which points defines the (new) triangle in this case?

e Given a large nonlinear optimization problem. Provide two reasons for choosing N-M compared to SQP, and two other reasons for *not* choosing N-M.

2 Optimality conditions and nonlinear problems (40%)

a Given the minimization problem (1).

What is the general form of f and c_i in the event of a QP-problem?

Specify a concrete example of a *convex* QP-problem where

$$n = 2, \quad \mathcal{E} = \{1\}, \quad \mathcal{I} = \{2, 3\}$$

Specify a concrete example of a *non-convex* QP-problem where

$$n = 2, \quad \mathcal{E} = \{1\}, \quad \mathcal{I} = \{2, 3\}$$

b Specify two conditions, related to f and the feasible region, which guarantee that (1) is a convex problem.

c The following theorem is stated for unconstrained problems, i.e. $\mathcal{E} = \varnothing, \quad \mathcal{I} = \varnothing$, (Textbook theorem 2.5)

When f is convex, any local minimizer is a global minimizer of f.

Prove this theorem.

d Given the minimization problem (1). Assume that

$$f = x_1 + \sqrt{3}x_2, \quad c_1(x) = x_1^2 + x_2^2 - 1, \quad \mathcal{E} = \{1\}, \quad \mathcal{I} = \varnothing$$

Show that the KKT-conditions are satisfied at $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$.

Check 2nd order conditions in both points to determine which point is (at least) a local minimum.

e Given the problem in **d)**. Further, assume that the optimization problem is solved using the SQP algorithm at the end of the Appendix.

Specify a suitable merit function ϕ_1.

The parameter μ_k should be included in the merit function above. What is its purpose?

Will the merit function generally to decrease from one iteration point to the next? Substantiate your answer.

Will the objective function f generally to decrease from one iteration point to the next? Substantiate your answer.

f Given the problem in **d)**. Further, assume that the optimization problem is solved using the SQP algorithm at the end of the Appendix.

Assume that the initial point x^0 is infeasible, $x^0 = (-2, 0)^T$, and that it converges to a point very close to $(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$ within 5 iterations. Hence, the iterates consist of the sequence $\{x^0, x^1, \ldots, x^5\}$. Sketch a possible sequence. Do you think the seqence will approach the feasible set immediately (x^1) or will this most likely happen after some interations?

3 Optimal control and MPC (30%)

a Assume the following problem. Note that the index relates to time.

$$\min \quad f_\infty = \frac{1}{2}\sum_{i=0}^{\infty}\{{x_i}^T Q x_i + {u_i}^T P u_i\}, \quad Q \succeq 0, \quad P \succ 0$$
$$x_{i+1} = Ax_i + Bu_i, \ 0 \le i \le \infty$$

The solution is given by the controller $u_i = Kx_i$. How is K computed? Specify the equations you need to apply for this purpose.

The controlled (closed-loop) system should be stable (meaning that $x_i \to 0$, $u_i \to 0$ when $i \to \infty$). Specify the conditions for this property. (There are two conditions).

What can be said about the eigenvalues of $A + BK$ when the closed-loop system is stable?

b A discrete nonlinear dynamic system is given in state space form

$$x_{k+1} = g(x_k, u_k), \quad x_k \in \mathbb{R}^3, \quad u_k \in \mathbb{R}^1, \quad u_k \in [0,1], \quad k = \{0, 1, \ldots\}$$

k is the time index, x_k are the states and u_k the control input (manipulated variable). Assume the system is controlled on a horizon $k \in \{0, 1, ..., N\}$, and that the initial state x_0 is known. Further, the deviation between the third element of the state vector (x_{3k}) and a time-varying reference value for this state (x_{3k}^{ref}) should be minimized.

Formulate a suitable optimization problem. Specify clearly which variables you are optimizing with respect to.

c Assume that it is important that the manipulated variable does not change too much from one time-step to the next time-step. How would you modify the objective function to allow for this?

d We apply the optimization problem specified in **b)** in MPC**.** Explain the MPC principle using a figure and state two important reasons for its industrial success.

e Assume that we include lower and upper bounds on the third state $\{\underline{x}_3, \overline{x}_3\}$. (To avoid any confusion we assume $\underline{x}_3 < x_{3k}^{ref} < \overline{x}_3$.) This may render the MPC optimization problem infeasible. Explain a method to avoid infeasible solutions and present a modified optimization problem.

Appendix

Part 1 Optimization problems and optimality conditions

$\mathcal{E}$ and $\mathcal{I}$ given below are two finite sets of indices.

General optimization problem. f and c_i are differentiable functions:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) \\ c_i(x) &= 0, \qquad i \in \mathcal{E} \\ c_i(x) &\geq 0, \qquad i \in \mathcal{I} \end{aligned} \tag{1}$$

The Lagrangian function is given by

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i c_i(x)$$

The KKT-conditions for (1) are given by:

$$\begin{aligned} \nabla_x \mathcal{L}(x^*,\lambda^*) &= 0 \\ c_i(x^*) &= 0, \qquad i \in \mathcal{E} \\ c_i(x^*) &\geq 0, \qquad i \in \mathcal{I} \\ \lambda_i^* &\geq 0, \qquad i \in \mathcal{I} \\ \lambda_i^* c_i(x^*) &= 0, \qquad i \in \mathcal{E}\cup\mathcal{I} \end{aligned} \tag{2}$$

2nd order (sufficient) conditions for (1) are given by:

$$w \in F_2(\lambda^*) \Leftrightarrow \begin{cases} \nabla c_i(x^*)^T w = 0 & \text{for all } i \in \mathcal{E} \\ \nabla c_i(x^*)^T w = 0 & \text{for all } i \in \mathcal{A}(x^*)\cap\mathcal{I} \text{ with } \lambda_i^* > 0 \\ \nabla c_i(x^*)^T w \geq 0 & \text{for all } i \in \mathcal{A}(x^*)\cap\mathcal{I} \text{ with } \lambda_i^* = 0 \end{cases}$$

Theorem (Second-Order Sufficient Conditions)
Suppose that for some feasible point $x^* \in R^n$ there is a Lagrange multiplier vector λ^* such that the KKT conditions (2) are satisfied. Suppose also that

$$w^T \nabla_{xx}\mathcal{L}(x^*,\lambda^*) w > 0, \qquad \text{for all } w \in F_2(\lambda^*),\ w \neq 0. \tag{3}$$

Then x^* is a strict local solution for (1).

LP-problem on standard form:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) = c^T x \\ s.t. \quad &Ax = b \\ &x \geq 0 \end{aligned}$$

where $A \in R^{m\times n}$ and $rank(A) = m$.

QP-problem on standard form:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) = \tfrac{1}{2}x^T G x + x^T d \\ s.t. \quad &a_i^T x = b_i, \qquad i \in \mathcal{E} \\ &a_i^T x \geq b_i, \qquad i \in \mathcal{I} \end{aligned}$$

where $G = G^T$. Alternatively, the equalities can be written $Ax = b$, $A \in \mathbb{R}^{m\times n}$.

Iterative method:

$$x_{k+1} = x_k + \alpha_k p_k$$
$$x_0 \ given$$
$$x_k, p_k \in \mathbb{R}^n, \ \alpha_k \in \mathbb{R}$$

p_k is the search direction and α_k is the line search parameter.

Part 2 Linear quadratic control of discrete dynamic systems

A typical optimal control problem on the time horizon 0 to n might take the form

$$\begin{aligned} \min \quad f_0 = & \frac{1}{2}\sum_{i=0}^{n-1}\{(y_i - y_{ref,i})^T Q_i (y_i - y_{ref,i}) \\ & + (u_i - u_{i-1})^T P_i (u_i - u_{i-1})\} \\ & + \frac{1}{2}(y_n - y_{ref,n})^T S (y_n - y_{ref,n}) \end{aligned} \tag{4}$$

subject to equality and inequality constraints

$$x_{i+1} = A_i x_i + B_i u_i, \; 0 \le i \le n-1 \tag{5}$$
$$y_i = H x_i$$
$$x_0 = \text{given (fixed)} \tag{6}$$
$$U_L \le u_i \le U_U, \; 0 \le i \le n-1 \tag{7}$$
$$Y_L \le y_i \le Y_U, \; 1 \le i \le n \tag{8}$$

where system dimensions are given by

$$\begin{aligned} u_i &\in \mathbb{R}^m \\ x_i &\in \mathbb{R}^l \\ y_i &\in \mathbb{R}^j \end{aligned}$$

The subscript i refers to the sampling instants. That is, subscript $i+1$ refers to the sample instant one sample interval after sample i. Note that the sampling time between each successive sampling instant is constant. Further, we assume that the control input u_i is constant between each sample.

Theorem: Assume that $x_{ref,i} = 0$, $u_{ref,i} = 0$, $0 \le i \le n$ and that $H = I$, i.e. $y_i = x_i$. The solution of (4), (5) and (6) is given by $u_i = K_i x_i$, $0 \le i \le n-1$ where the feedback gain matrix is derived by

$$\begin{aligned} K_i &= -P_i^{-1} B_i^T R_{i+1} (I + B_i P_i^{-1} B_i^T R_{i+1})^{-1} A_i, \; 0 \le i \le n-1 \\ R_i &= Q_i + A_i^T R_{i+1} (I + B_i P_i^{-1} B_i^T R_{i+1})^{-1} A_i, \; 0 \le i \le n-1 \\ R_n &= S \end{aligned}$$

18.4 A PRACTICAL LINE SEARCH SQP METHOD

From the discussion in the previous section, we can see that there is a wide variety of line search SQP methods that differ in the way the Hessian approximation is computed, in the step acceptance mechanism, and in other algorithmic features. We now incorporate some of these ideas into a concrete, practical SQP algorithm for solving the nonlinear programming problem (18.10). To keep the description simple, we will not include a mechanism such as (18.12) to ensure the feasibility of the subproblem, or a second-order correction step. Rather, the search direction is obtained simply by solving the subproblem (18.11). We also assume that the quadratic program (18.11) is convex, so that we can solve it by means of the active-set method for quadratic programming (Algorithm 16.3) described in Chapter 16.

Algorithm 18.3 (Line Search SQP Algorithm).

Choose parameters $\eta \in (0, 0.5)$, $\tau \in (0, 1)$, and an initial pair (x_0, λ_0);
Evaluate $f_0, \nabla f_0, c_0, A_0$;
If a quasi-Newton approximation is used, choose an initial $n \times n$ symmetric positive definite Hessian approximation B_0, otherwise compute $\nabla^2_{xx}\mathcal{L}_0$;
repeat until a convergence test is satisfied
    Compute p_k by solving (18.11); let $\hat{\lambda}$ be the corresponding multiplier;
    Set $p_\lambda \leftarrow \hat{\lambda} - \lambda_k$;
    Choose μ_k to satisfy (18.36) with $\sigma = 1$;
    Set $\alpha_k \leftarrow 1$;
    while $\phi_1(x_k + \alpha_k p_k; \mu_k) > \phi_1(x_k; \mu_k) + \eta\alpha_k D_1(\phi(x_k; \mu_k)p_k)$
        Reset $\alpha_k \leftarrow \tau_\alpha \alpha_k$ for some $\tau_\alpha \in (0, \tau]$;
    end (while)
    Set $x_{k+1} \leftarrow x_k + \alpha_k p_k$ and $\lambda_{k+1} \leftarrow \lambda_k + \alpha_k p_\lambda$;
    Evaluate $f_{k+1}, \nabla f_{k+1}, c_{k+1}, A_{k+1}$, (and possibly $\nabla^2_{xx}\mathcal{L}_{k+1}$);
    If a quasi-Newton approximation is used, set
        $s_k \leftarrow \alpha_k p_k$ and $y_k \leftarrow \nabla_x \mathcal{L}(x_{k+1}, \lambda_{k+1}) - \nabla_x \mathcal{L}(x_k, \lambda_{k+1})$,
    and obtain B_{k+1} by updating B_k using a quasi-Newton formula;
end (repeat)

We can achieve significant savings in the solution of the quadratic subproblem by warm-start procedures. For example, we can initialize the working set for each QP subproblem to be the final active set from the previous SQP iteration.

We have not given particulars of the quasi-Newton approximation in Algorithm 18.3. We could use, for example, a limited-memory BFGS approach that is suitable for large-scale problems. If we use an exact Hessian $\nabla^2_{xx}\mathcal{L}_k$, we assume that it is modified as necessary to be positive definite on the null space of the equality constraints.

Instead of a merit function, we could employ a filter (see Section 15.4) in the inner "while" loop to determine the steplength α_k. As discussed in Section 15.4, a feasibility restoration phase is invoked if a trial steplength generated by the backtracking line search is