Department of Engineering Cybernetics
Faculty of Information Technology, Mathematics and Electrical Engineering
Norwegian University of Science and Technology (NTNU)

Contact for questions during exam:
Postdoc Mansoureh Jesmani
Tel: 47951195

English version

Exam in TTK4135

Optimization and Control

Optimalisering og regulering

Saturday May 28, 2016

Time: 09:00 – 13:00

Combination of allowed help remedies:
D — No printed or hand-written notes.
Certified calculator with empty memory.
Grading date: June 25

In the Appendix potentially useful information is included.

1 Linear programming (LP) (40 %)

a (12 %) We are interested in maximizing the profit for a farmer who grows apples and bananas. The following information is provided:

- A farmer wants to grow apples A (use x_1 for apples) and bananas B (use x_2 for bananas).
- He has a field of size 100 000 m^2.
- Growing 1 tonne of A requires an area of 4 000 m^2,
- Growing 1 tonne of B requires an area of 3 000 m^2.
- A requires 60 kg fertilizer per tonne grown,
- B requires 80 kg fertilizer per tonne grown.

- The profit for A is 7000 per tonne (includes fertilizer cost).
- The profit for B is 6000 per tonne (includes fertilizer cost).
- The farmer can legally use up to 2000 kg of fertilizer.
- Maximize profit.

Describe the problem on standard form (as in the Appendix (A.6)). Use m^2, kg and tonne as units for area, fertilizer and apples/bananas, respectively. Specify all matrices and vectors.

b (4 %) The Simplex method is appropriate for solving the apples and bananas problem above. Does the Simplex method use gradients in its search for the solution? Does the Simplex algorithm require a feasible starting point?

c (3 %) Is the apples and bananas problem as described above a convex problem? Justify your answer.

d (7 %) A Simplex algorithm needs to solve many linear equations. Usually LU factorization is used for this purpose. Assume that we need to solve $Cz = d$ and that an LU factorization of C is available.

Questions:

- What is meant by LU factorization and what is the structure of the factorized matrices?
- Explain how $Cz = d$ is solved when C is factorized?

e (8 %) Assume that we change the problem in a) by including changing profits.

- The profit for A is $7000 - 200x_1$ per tonne.
- The profit for B is $6000 - 140x_2$ per tonne.

All other information is kept unchanged.

Questions:

- Formulate the new problem.
- Which type of optimization problem is this?
- Is the new problem a convex problem?

f (6 %) The solution in a) is $x_1^* = 14.3$ and $x_2^* = 14.3$ with a profit of 185714. The optimal Lagrange multiplier for the area constraint is $\lambda_1^* = 1.4$ and for the fertilizer constraint is $\lambda_2^* = 21.4$. What will the profit be if we increase the legal use of fertilizer from 2000 kg to 2001 kg? Justify your answer.

2 MPC and optimal control (40 %)

a (4 %) There exist two important classes of methods for optimizing dynamic systems.

- Quasi dynamic optimization: Optimize a dynamic system by repetitive optimization on a static model.
- Dynamic optimization: Optimize on a dynamic model. In this case the solution will be a function of time, i.e., all decision variables will be functions of time.

Discuss the advantages and drawbacks of using 'Quasi dynamic optimization' compared to 'Dynamic optimization'.

b (26 %) The formulation below has been used several times in the course.

$$\min_{z \in \mathbb{R}^n} f(z) = \sum_{t=0}^{N-1} \frac{1}{2} x_{t+1}^\top Q_{t+1} x_{t+1} + d_{xt+1} x_{t+1} + \frac{1}{2} u_t^\top R_t u_t + d_{ut} u_t \tag{1a}$$

subject to

$$x_{t+1} = A_t x_t + B_t u_t, \qquad t = 0, \dots, N-1 \tag{1b}$$

$$x_0,\ u_{-1} = \text{given} \tag{1c}$$

$$x^{\text{low}} \le x_t \le x^{\text{high}}, \qquad t = 1, \dots, N \tag{1d}$$

$$u^{\text{low}} \le u_t \le u^{\text{high}}, \qquad t = 0, \dots, N-1 \tag{1e}$$

$$-\Delta u^{\text{high}} \le \Delta u_t \le \Delta u^{\text{high}}, \qquad t = 0, \dots, N-1 \tag{1f}$$

where

$$Q_t \succeq 0 \qquad t = 1, \dots, N \tag{1g}$$

$$R_t \succeq 0 \qquad t = 0, \dots, N-1 \tag{1h}$$

$$\Delta u_t = u_t - u_{t-1} \tag{1i}$$

$$z^\top = (x_1^\top, \dots, x_N^\top, u_0^\top \dots, u_{N-1}^\top) \tag{1j}$$

$$n = N \cdot (n_x + n_u) \tag{1k}$$

Questions:

- b1 Why is the problem called an open loop optimization problem?
- b2 Reformulate the dynamic model as an LTI (linear time invariant) model instead of an LTV (linear time varying) model.
- b3 Why is (1f) usually included in practical formulations?
- b4 Which type of optimization problem is this?
- b5 Assume that the dynamic model is nonlinear instead of linear. Which type of optimization problem would you then have, and suggest a suitable solution algorithm.
- b6 The optimization problem may be infeasible. Suggest a reformulation that guarantees feasibility.
- b7 Assume that the prediction horizon is 12 steps, that there are 5 states (in the state vector), and that there are 2 control inputs. How many optimization variables is there in the problem above (this is often called a full space formulation)? How many optimization variables are there in a reduced space formulation of the problem above? Discuss briefly the pros and cons of using the full space problem vs. the reduced space problem.

c (10 %) Explain the MPC principle through an algorithm (use macro code with a resolution similar to the algorithms specified in the lecture notes). Further, explain MPC with a figure with time along the horizontal axis.

3 The Rosenbrock function (20 %)

a (8 %) We now focus on the Rosenbrock function

$$\begin{aligned} f(x_1, x_2) &= (a - x_1)^2 + b(x_2 - x_1^2)^2 \\ a, b &\geq 0 \end{aligned} \tag{2}$$

which has its global minimum at (a, a^2).

Compute the gradient $\nabla f(x_1, x_2)$ and the Hessian matrix $\nabla^2 f(x_1, x_2)$.

b (6 %) Select $a = 1$ and $b = 2$. Show that the first and second order conditions are satisfied at the solution.

c (6 %) Formulate an optimization problem where the Rosenbrock function is minimized subject to positivity constraints (x_1 and x_2 should be positive). Is the feasible set a convex set? Is the optimization problem a convex problem?

Appendix

Part 1 Optimization Problems and Optimality Conditions

A general formulation for constrained optimization problems is

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) && \text{(A.1a)}\\ \text{s.t.} \quad & c_i(x) = 0, \qquad i \in \mathcal{E} && \text{(A.1b)}\\ & c_i(x) \geq 0, \qquad i \in \mathcal{I} && \text{(A.1c)} \end{aligned}$$

where f and the functions c_i are all smooth, differentiable, real-valued functions on a subset of $\mathbb{R}^n$, and $\mathcal{E}$ and $\mathcal{I}$ are two finite sets of indices.

The Lagrangean function for the general problem (A.1) is

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i c_i(x) \tag{A.2}$$

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

$$\begin{aligned} \nabla_x \mathcal{L}(x^*,\lambda^*) &= 0 && \text{(A.3a)}\\ c_i(x^*) &= 0, \qquad i \in \mathcal{E} && \text{(A.3b)}\\ c_i(x^*) &\geq 0, \qquad i \in \mathcal{I} && \text{(A.3c)}\\ \lambda_i^* &\geq 0, \qquad i \in \mathcal{I} && \text{(A.3d)}\\ \lambda_i^* c_i(x^*) &= 0, \qquad i \in \mathcal{E}\cup\mathcal{I} && \text{(A.3e)} \end{aligned}$$

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

$$w \in \mathcal{C}(x^*,\lambda^*) \Leftrightarrow \begin{cases} \nabla c_i(x^*)^\top w = 0 & \text{for all } i \in \mathcal{E} \\ \nabla c_i(x^*)^\top w = 0 & \text{for all } i \in \mathcal{A}(x^*)\cap\mathcal{I} \text{ with } \lambda_i^* > 0 \\ \nabla c_i(x^*)^\top w \geq 0 & \text{for all } i \in \mathcal{A}(x^*)\cap\mathcal{I} \text{ with } \lambda_i^* = 0 \end{cases} \tag{A.4}$$

Theorem 1: (Second-Order Sufficient Conditions) *Suppose that for some feasible point* $x^* \in \mathbb{R}^n$ *there is a Lagrange multiplier vector* λ^* *such that the KKT conditions* (A.3) *are satisfied. Suppose also that*

$$w^\top \nabla^2_{xx}\mathcal{L}(x^*,\lambda^*) w > 0, \qquad \text{for all } w \in \mathcal{C}(x^*,\lambda^*),\ w \neq 0. \tag{A.5}$$

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

LP problem in standard form:

$$\begin{aligned} \min_x \quad & f(x) = c^\top x && \text{(A.6a)} \\ \text{s.t.} \quad & Ax = b && \text{(A.6b)} \\ & x \geq 0 && \text{(A.6c)} \end{aligned}$$

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

QP problem in standard form:

$$\begin{aligned} \min_x \quad & f(x) = \frac{1}{2} x^\top G x + x^\top c && \text{(A.7a)} \\ \text{s.t.} \quad & a_i^\top x = b_i, \qquad i \in \mathcal{E} && \text{(A.7b)} \\ & a_i^\top x \geq b_i, \qquad i \in \mathcal{I} && \text{(A.7c)} \end{aligned}$$

where G is a symmetric $n \times n$ matrix, $\mathcal{E}$ and $\mathcal{I}$ are finite sets of indices and c, x and $\{a_i\}, i \in \mathcal{E} \cup \mathcal{I}$, are vectors in $\mathbb{R}^n$. Alternatively, the equalities can be written $Ax = b$, $A \in \mathbb{R}^{m \times n}$.

Iterative method:

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k p_k && \text{(A.8a)} \\ & x_0 \text{ given} && \text{(A.8b)} \\ x_k, p_k &\in \mathbb{R}^n, \ \alpha_k \in \mathbb{R} && \text{(A.8c)} \end{aligned}$$

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

Part 2 Optimal Control

A typical open-loop optimal control problem on the time horizon 0 to N is

$$\min_{z \in \mathbb{R}^n} f(z) = \sum_{t=0}^{N-1} \frac{1}{2} x_{t+1}^\top Q_{t+1} x_{t+1} + d_{xt+1} x_{t+1} + \frac{1}{2} u_t^\top R_t u_t + d_{ut} u_t \tag{A.9a}$$

subject to

$$x_{t+1} = A_t x_t + B_t u_t, \qquad t = 0, \ldots, N-1 \tag{A.9b}$$
$$x_0 = \text{given} \tag{A.9c}$$
$$x^{\text{low}} \leq x_t \leq x^{\text{high}}, \qquad t = 1, \ldots, N \tag{A.9d}$$
$$u^{\text{low}} \leq u_t \leq u^{\text{high}}, \qquad t = 0, \ldots, N-1 \tag{A.9e}$$
$$-\Delta u^{\text{high}} \leq \Delta u_t \leq \Delta u^{\text{high}}, \qquad t = 0, \ldots, N-1 \tag{A.9f}$$
$$Q_t \succeq 0 \qquad t = 1, \ldots, N \tag{A.9g}$$
$$R_t \succeq 0 \qquad t = 0, \ldots, N-1 \tag{A.9h}$$

where

$$u_t \in \mathbb{R}^{n_u} \tag{A.9i}$$
$$x_t \in \mathbb{R}^{n_x} \tag{A.9j}$$
$$\Delta u_t = u_t - u_{t-1} \tag{A.9k}$$
$$z^\top = (x_1^\top, \ldots, x_N^\top, u_0^\top, \ldots, u_{N-1}^\top) \tag{A.9l}$$

The subscript t denotes discrete time sampling instants.

The optimization problem for linear quadratic control of discrete dynamic systems is given by

$$\min_{z \in \mathbb{R}^n} f(z) = \sum_{t=0}^{N-1} \frac{1}{2} x_{t+1}^\top Q_{t+1} x_{t+1} + \frac{1}{2} u_t^\top R_t u_t \tag{A.10a}$$

subject to

$$x_{t+1} = A_t x_t + B_t u_t \tag{A.10b}$$
$$x_0 = \text{given} \tag{A.10c}$$

where

$$u_t \in \mathbb{R}^{n_u} \tag{A.10d}$$
$$x_t \in \mathbb{R}^{n_x} \tag{A.10e}$$
$$z^\top = (x_1^\top, \ldots, x_N^\top, u_0^\top \ldots, u_{N-1}^\top) \tag{A.10f}$$

Theorem 2: *The solution of (A.10) with $Q_t \succeq 0$ and $R_t \succ 0$ is given by*

$$u_t = -K_t x_t \tag{A.11a}$$

where the feedback gain matrix is derived by

$$K_t = R_t^{-1} B_t^\top P_{t+1}(I + B_t R_t^{-1} B_t^\top P_{t+1})^{-1} A_t, \qquad t = 0, \dots, N-1 \tag{A.11b}$$

$$P_t = Q_t + A_t^\top P_{t+1}(I + B_t R_t^{-1} B_t^\top P_{t+1})^{-1} A_t, \qquad t = 0, \dots, N-1 \tag{A.11c}$$

$$P_N = Q_N \tag{A.11d}$$

Part 3 Sequential quadratic programming (SQP)

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)