Department of Engineering Cybernetics

Examination paper for TTK4135 Optimization and Control

Academic contact during examination: PhD student Anders Lindahl Flåten

Phone: 92869057

Examination date: Saturday 20.5.2017

Examination time (from-to): 09:00-13:00

Permitted examination support material: D – No printed or hand-written notes – Certified calculator with empty memory

Other information:

Language: English

Number of pages (front page excluded): 10

Number of pages enclosed: 11

Informasjon om trykking av eksamensoppgave
Originalen er:
1-sidig X 2-sidig □
sort/hvit X farger □
skal ha flervalgskjema □

Checked by:

Date Signature

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

Contact for questions during exam:
PhD student Andreas Lindahl Flåten
Tel: 92869057

English version

Exam in TTK4135

Optimization and Control

Optimalisering og regulering

Saturday May 20, 2017

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 14

In the Appendix potentially useful information is included.

1 Quadratic programming (QP) (34 %)

Below we formulate a QP-problem

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

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

a (6 %) Formulate the KKT conditions for the QP problem above.

b (6 %) We now focus on the complementarity condition, which is part of the KKT conditions from the above question.

- What does strict complimentarity mean?
- Why is strict complimentarity desirable?

c (4 %) Assume $\mathcal{E} = \emptyset$, $\mathcal{I} = \emptyset$ (i.e. no equality/inequality constraints) and $G \succ 0$.

- Show that the Newton direction is given by $p = -G^{-1}(Gx + d)$.
- Show that p is a descent direction when $p \neq 0$.

d (6 %) Assume $\mathcal{E} = \{1, 2, 3\}$, $\mathcal{I} = \emptyset$ and $n = 4$

Suggest two A matrices; one where the LICQ condition holds and one where the LICQ does not hold.

e (12 %) Formulate the following investment problem as a QP-problem (you shall not solve the QP-problem).

- A corporation requires 60 million kroner to finance a new manufacturing process.
- Three different banks have agreed to supply all or parts of this amount.
- All three banks insist that the loan plus interest charges be repaid over 6 years.
- The repayment schedules differ from bank to bank, as shown in the table below. The table should be interpreted as follows:
  - During year 1 Bank 1 requires no repay, Bank 2 requires a repay of 5% of the loan, and Bank 3 requires a repay of 40% of the loan.
  - Bank 1 should during 6 years be repaid by 175% a of the loan (175% = 0% + 0% + 30% + 40% + 50% + 55%).
- To simplify we assume zero discount factor meaning that the value of 1 krone today is the same as the value of 1 krone next year.

	Percent (%) of amount to be repaid each year					
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Bank 1	0	0	30	40	50	55
Bank 2	5	15	25	35	40	45
Bank 3	40	40	0	35	15	15

The corporation believes it is advantageous to borrow such that the total yearly payments on the loan are as equal as possible, yet it does not wish to pay more than 40 million kroner in total interest charges. Formulate a QP-problem to find the amount borrowed from each bank to satisfy the corporation's goal and constraints.

(Hint: It may be beneficial to define the decision variables first, and also to define an extra variable equal to the average yearly payment).

2 KKT conditions (12 %)

a (6 %)

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & 2x_1 + x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 4 \\ & x_2 \geq 0 \end{aligned}$$

Show that the solution is $x^* = (-2.0, 0.0)^T$, and find the Lagrange multipliers.

b (6 %) Add the constraint $x_1 \geq 1.0$. Find the solution and Lagrange multipliers.

3 MPC and optimal control (24 %)

a (6 %) Change the formulation in (A.9) to include a time-invariant objective function and a linear time-invariant model.

b (12 %) An LQ-controller is obtainted by removing the inequality constraints in (A.9). A formulation is shown in (A.10) and the corresponding control solution is given in (A.11).
Assume that A_t, B_t, Q_t, R_t are time-invariant matrices, i.e. $A_t = A$, $B_t = B$, $Q_t = Q$, $R_t = R$, and that the prediction horizon extends to infinity ($N \to \infty$). Derive the control solution for this infinite horizon LQ-problem (Hint: The solution can be derived from (A.11)).

c (6 %) Explain the main advantages and disadvantages of using a linear MPC controller.

4 Sequential quadratic programming (SQP) (30 %)

We now focus our attention on the SQP-algorithm. A pseudo code for a SQP-algorithm is shown in Algorithm 18.3 in the Appendix.

a (6 %) We focus on the line search part of Algorithm 18.3, i.e. computation of α_k. Explain why this part of the algorithm satisfies the first Wolfe condition (also called Armijo condition). Illustrate this condition using a figure.

b (6 %) Assume that the line search part of the algorithm uses the l_1 merit function shown below (we use the same notation as in the course textbook).

$$\phi_1(x;\mu) = f(x) + \mu \sum_{i \in E} |c_i(x)| + \mu \sum_{i \in I} |c_i(x)|^- \tag{2}$$

Why is it necessary to perform the line search on a merit function instead of the objective function itself? Please include a simple graphical sketch to support the explanation.

c (6 %) Define the function $|c_i(x)|^-$ in (2). What are the values of $|3.5|^-$ and $|-2.7|^-$, respectively?

d (6 %) The parameter μ in (2) will in an SQP-algorithm vary from one iteration to the next. Will it typically increase with increasing iteration number k or decrease with increasing iteration number k? Please justify your answer.

e (6 %) Assume that Algorithm 18.3 is used to solve a non-convex problem. If the algorithm converges to a point x' satisfying local conditions, we know that x' is a local solution. Assume that it is important to get close to the global solution. Suggest a method to obtain a solution close to the global solution.

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

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

where

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

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

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

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, ∇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}, ∇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)