NTNU – Trondheim
Norwegian University of
Science and Technology

For questions during the exam:
Lars Imsland, tel. 47 23 19 49.

Exam in TTK4135 Optimization and Control

Monday, May 28th 2018

09:00 – 13:00

Permitted aids (code D): No printed or hand-written support material is allowed. A specific basic calculator is allowed.

Answers in English, Norwegian, or a mixture of the two accepted.

Grades available: As specified by regulations.

Problem 1 (25 %)

Consider the following optimization problem:

$$\begin{aligned} \min_{x} \quad & 4x_1^2 - x_1x_2 + 6x_2^2 + 2x_3^2 + 10x_1 - 2x_2 - 5x_3 \\ \text{subject to} \quad & x_2 + x_3 = 2 \\ & x_2 = \theta \end{aligned}$$

where θ is an unknown constant.

(5 %) (a) Is this is a convex optimization problem? Give reason for your answer.

Solution: The objective can be written

$$\frac{1}{2}x^\top Qx + c^\top x$$

where

$$Q = \begin{pmatrix} 8 & -1 & 0 \\ -1 & 12 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{and} \quad c^\top = \begin{pmatrix} 10 & -2 & -5 \end{pmatrix}.$$

Since Q is positive definite (all leading minors are positive), and the constraints are linear (for all constant θ), the optimization problem is convex. Q positive definite: 3p, constraints linear: 2p.

(2 %) (b) What type of optimization problem is this?

Solution: This is a quadratic program (QP).

(8 %) (c) Write down the KKT conditions for this problem. The answer should contain numbers.

Solution: Write the constraints as $Ax = b$, with

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ \theta \end{pmatrix}$$

The Lagrangian $\mathcal{L}(x, \lambda) = \frac{1}{2}x^\top Q x + c^\top x - \lambda^\top (Ax - b)$, and the KKT conditions are

$$\begin{pmatrix} Q & -A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}$$

$$\begin{pmatrix} 8 & -1 & 0 & 0 & 0 \\ -1 & 12 & 0 & -1 & -1 \\ 0 & 0 & 4 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} -10 \\ 2 \\ 5 \\ 2 \\ \theta \end{pmatrix}.$$

The answer need not be on matrix form for full score, to list KKT conditions with numbers inserted is sufficient. Only stationarity conditions ($Qx + c - A^\mathsf{T}\lambda = 0$) gives 4p.

(6 %) (d) What is the solution x^* to the optimization problem as a function of θ?

Solution: Since the problem is convex, the KKT conditions are necessary and sufficient for a solution. From the last line of the KKT conditions, we see that $x_2 = \theta$. Inserting this in the line above, $x_3 = 2 - \theta$. The first line is $8x_1 - x_2 = -10$, rearranging to $x_1 = \frac{x_2}{8} - \frac{10}{8} = \frac{\theta}{8} - \frac{5}{4}$. That is,

$$x^* = \begin{pmatrix} \frac{1}{8} \\ 1 \\ -1 \end{pmatrix} \theta + \begin{pmatrix} -\frac{5}{4} \\ 0 \\ 2 \end{pmatrix}.$$

Correct reasoning but minor mistakes give 2-4 points.

(Note that this can also be solved by substituting the equality constraints to obtain an optimization problem in one variable. However, the next problem then becomes somewhat more difficult.)

(4 %) (e) What are the optimal Lagrangian multipliers λ^* for the constraints, as a function of θ?

Solution: From the third line of the KKT conditions, $4x_3 - \lambda_1 = 5$ giving $\lambda_1 = 4x_3 - 5 = 3 - 4\theta$. The second line is $-x_1 + 12x_2 - \lambda_1 - \lambda_2 = 2$, giving $\lambda_2 = -2 - (\frac{\theta}{8} - \frac{5}{4}) + 12\theta - (3 - 4\theta) = -15\frac{7}{8}\theta + 3\frac{3}{4}$.

$$\lambda^* = \begin{pmatrix} -4 \\ 15\frac{7}{8} \end{pmatrix} \theta + \begin{pmatrix} 3 \\ -3\frac{3}{4} \end{pmatrix}.$$

(That is, the solution is affine (linear) in x. For a general QP, the solution (and Lagrange multipliers) will be piecewise affine in the elements of b.)

Problem 2 (20 %)

Emma is starting a micro brewery, and needs to plan her efforts.

- She is able to produce and sell maximum 100 boxes of beer each day.
- She can work up to 14 hours per day.
- It takes her 1 hour to produce 10 boxes of light beer.
- It takes her 2 hours to produce 10 boxes of dark beer.
- She earns 20 Euros for one box of light beer.
- She earns 30 Euros for one box of dark beer.

She wants to mazimize her profits.

(10 %) (a) Formulate the optimization problem. What type of optimization problem is this?

Solution: Let x_1 be number of boxes of light beer, and x_2 be numer of boxes of dark beer.

$$\begin{aligned}
\max_{x_1,\, x_2} \quad & 20x_1 + 30x_2 \\
\text{subject to:} & x_1 + x_2 \leq 100 \\
& 0.1x_1 + 0.2x_2 \leq 14 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}$$

A Linear program (LP).
Some might use x_1, x_2 as number of 10 boxes. This is possible, provided one change the coefficients, for example:

$$\begin{aligned}
\max_{x_1,\, x_2} \quad & 200x_1 + 300x_2 \\
\text{subject to:} & x_1 + x_2 \leq 10 \\
& x_1 + 2x_2 \leq 14 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}$$

(10 %) (b) Draw/sketch the feasible region, and the contours of the objective function. What is the optimal solution? It is enough to indicate the optimal solution in the figure, you need not find the exact numbers.

Solution:

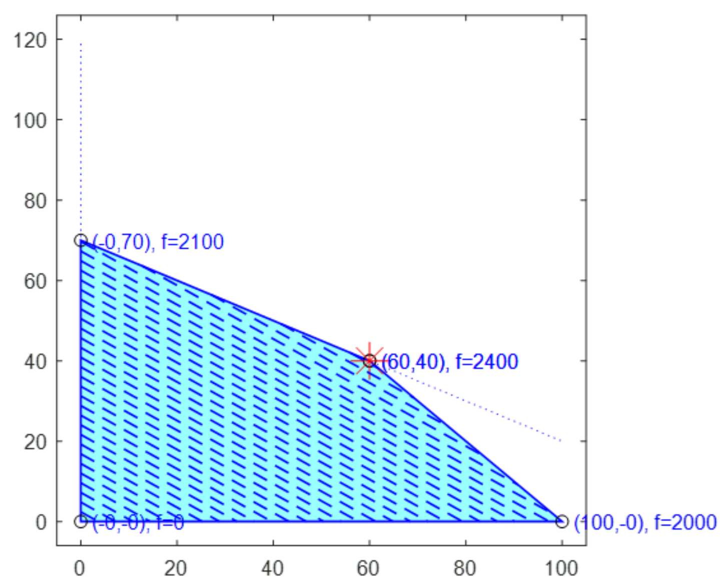

Figure 1: Feasible region.

Since it is an LP, we can check all the "corners", and see that it is optimal to produce 60 boxes of light beer and 40 boxes of dark beer. Alternatively, we see from the contours of the objective function which corner must be optimal.

Problem 3 (26 %)

Consider the optimization problem

$$\min_{z \in \mathbb{R}^n} \sum_{j=1}^{n} c_j |z_j| \quad \text{subject to} \quad Az \geq b \tag{1}$$

where $|\cdot|$ is the absolute value, and all c_j are positive.

(4 %) (a) Do the KKT conditions apply to this optimization problem? Explain.

> **Solution:** Since the objective function is not differentiable (smooth), the KKT conditions does not apply.

The optimization problem

$$\min_{z^+, z^- \in \mathbb{R}^n} \sum_{j=1}^{n} c_j (z_j^+ + z_j^-) \quad \text{subject to} \quad A(z^+ - z^-) \geq b, \quad z_j^+ \geq 0, \quad z_j^- \geq 0 \tag{2}$$

is equivalent to (1).

(10 %) (b) Explain ("prove") why these two optimization problems are equivalent. Hint: Let $z_j = z_j^+ - z_j^-$ and $|z_j| = z_j^+ + z_j^-$.

> **Solution:** The optimal solutions of both optimization problems are the same if, for each j, at least one of the values z_j^+ and z_j^- is zero. In that case, $z_j = z_j^+$ when $z_j \geq 0$, and $z_j = -z_j^-$ when $z_j \leq 0$. Assume for a moment that the optimal values of z_j^+ and z_j^- are both positive for a particular j, and let $\delta = \min\{z_j^+, z_j^-\}$. Subtracting $\delta > 0$ from both z_j^+ and z_j^- leaves the value of $z_j = z_j^+ - z_j^-$ unchanged, but reduces the value of $|z_j| = z_j^+ + z_j^-$ by 2δ. This contradicts the optimality assumption, because the objective function value can be reduced by $2\delta c_j$.
>
> This explanation is copied from the *AIMMS Modeling Guide - Linear Programming Tricks*. Other explanations are possible, but they should include the necessity of positive c_j, and that z_j^+ is the positive part of z_j, and z_j^- the negative part, and only one of them can be positive at the same time.
>
> Only copying the hint into (2) to obtain (1) is 2p. Recognizing that z_j^+ is the positive part of z_j, and z_j^- the negative part: 2p. Realizing that due to positive c_j only one can be positive at the same time: full score.

(4 %) (c) Do the KKT conditions apply to (2)? What type of optimization problem is this?

> **Solution:** Yes. It is an LP.

(8 %) (d) Use the above to formulate the open loop dynamic optimization problem

$$\begin{aligned}
\min_{u_0, u_1, x_1, x_2} \quad & \sum_{i=0}^{1} q_i |x_{i+1}| + r_i |u_i| \\
\text{subject to} \quad & x_{i+1} = 1.2 x_i + u_i, \quad i = 0, 1 \\
& |x_i| \leq 1, \quad i = 1, 2 \\
& |u_i| \leq 1, \quad i = 0, 1
\end{aligned}$$

as a linear program (LP). Here, q_i and r_i are positive constants.

Solution: (The problem could have mentioned that x_0 is given. Assuming it is a variable is OK.

Correct answers should introduce two new variables for each of u_0, u_1, x_1 and x_2 (and possibly x_0), and insert into the problem above.

Problem 4 (15 %)

Given the nonlinear programming problem

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \quad c(x) = 0, \tag{3}$$

where $c(x)^\top = [c_1(x), c_2(x), \ldots, c_m(x)]$. Define

$$F(x,\lambda) = \begin{bmatrix} \nabla f(x) - A(x)^\top \lambda \\ c(x) \end{bmatrix} = 0, \tag{4}$$

where $A(x) = \begin{bmatrix} \nabla c_1(x) & \nabla c_2(x) & \ldots & \nabla c_m(x) \end{bmatrix}^\top$.

(10 %) (a) Formulate Newton's method, in simple pseude-code, for solving $F(x,\lambda) = 0$.

Solution: See book, Algorithm 11.1. Students that use Newton's method from Ch. 3, get 2p (that is, defining the Newton direction).

(5 %) (b) Explain why, and to what extent, the algorithm in (a) solves (3) (can be answered also if you did not answer (a)).

Solution: Since $F(x,\lambda) = 0$ are the KKT conditions for (3), the algorithm in (a) (if it converges) provides a point that fulfills the necessary conditions, that is, it is a candidate solution to (3).

Full score for KKT and "necessary condition".

Problem 5 (14 %)

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

$$\min_{z} f^{\infty}(z) = \sum_{t=0}^{\infty} \frac{1}{2} x_{t+1}^{\top} Q x_{t+1} + \frac{1}{2} u_t^{\top} R u_t \tag{5a}$$

subject to

$$x_{t+1} = A x_t + B u_t \tag{5b}$$
$$x_0 = \text{given} \tag{5c}$$

where

$$u_t \in \mathbb{R}^{n_u} \tag{5d}$$
$$x_t \in \mathbb{R}^{n_x} \tag{5e}$$
$$z^{\top} = (u_0^{\top}, x_1^{\top}, u_1^{\top}, x_2^{\top}, \ldots) \tag{5f}$$

(8 %) (a) Assume (A, B) stabilizable. Write down the equations defining the infinite horizon LQ controller. Hint: Theorem 2 in the back may be useful.

Solution: The solution is the steady-state (static) solution of the Riccati equations in Theorem 2. When these converge ($N \to \infty$), a static solution must fulfill (A11b) -(A11c), that is we can remove the indices in (A11):

$$u_t = -K x_t \tag{6a}$$

where the feedback gain matrix K is derived by

$$K = R^{-1} B^{\top} P (I + B R^{-1} B^{\top} P)^{-1} A \tag{6b}$$
$$P = Q + A^{\top} P (I + B R^{-1} B^{\top} P)^{-1} A \tag{6c}$$

(6 %) (b) Draw a block diagram of the output feedback LQ controller (the LQG controller).

Solution: See Figure 4.7 in F&H.

Appendix

Part 1 Optimization Problems and Optimality Conditions

A general formulation for constrained optimization problems is

$$\min_{x\in\mathbb{R}^n} \quad f(x) \tag{A1a}$$
$$\text{s.t.} \quad c_i(x) = 0, \qquad i \in \mathcal{E} \tag{A1b}$$
$$c_i(x) \geq 0, \qquad i \in \mathcal{I} \tag{A1c}$$

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 Lagrangian function for the general problem (A1) is

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i c_i(x) \tag{A2}$$

The KKT-conditions for (A1) are given by:

$$\nabla_x \mathcal{L}(x^*,\lambda^*) = 0 \tag{A3a}$$
$$c_i(x^*) = 0, \qquad i \in \mathcal{E} \tag{A3b}$$
$$c_i(x^*) \geq 0, \qquad i \in \mathcal{I} \tag{A3c}$$
$$\lambda_i^* \geq 0, \qquad i \in \mathcal{I} \tag{A3d}$$
$$\lambda_i^* c_i(x^*) = 0, \qquad i \in \mathcal{E}\cup\mathcal{I} \tag{A3e}$$

2nd order (sufficient) conditions for (A1) are given by:

$$w \in \mathcal{C}(x^*,\lambda^*) \Leftrightarrow \begin{cases} \nabla c_i(x^*)^\top w = 0 & \text{for all } i \in \mathcal{E} \\ \nabla c_i(x^*)^\top w = 0 & \text{for all } i \in \mathcal{A}(x^*)\cap\mathcal{I} \text{ with } \lambda_i^* > 0 \\ \nabla c_i(x^*)^\top w \geq 0 & \text{for all } i \in \mathcal{A}(x^*)\cap\mathcal{I} \text{ with } \lambda_i^* = 0 \end{cases} \tag{A4}$$

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* (A3) *are satisfied. Suppose also that*

$$w^\top \nabla^2_{xx}\mathcal{L}(x^*,\lambda^*) w > 0, \qquad \text{for all } w \in \mathcal{C}(x^*,\lambda^*),\ w \neq 0. \tag{A5}$$

Then x^* is a strict local solution for (A1).

LP problem in standard form:

$$\min_{x} \quad f(x) = c^\top x \tag{A6a}$$
$$\text{s.t.} \quad Ax = b \tag{A6b}$$
$$x \geq 0 \tag{A6c}$$

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

QP problem in standard form:

$$\min_{x} \quad f(x) = \frac{1}{2} x^\top G x + x^\top c \tag{A7a}$$
$$\text{s.t.} \quad a_i^\top x = b_i, \qquad i \in \mathcal{E} \tag{A7b}$$
$$a_i^\top x \geq b_i, \qquad i \in \mathcal{I} \tag{A7c}$$

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:

$$x_{k+1} = x_k + \alpha_k p_k \tag{A8a}$$
$$x_0 \text{ given} \tag{A8b}$$
$$x_k, p_k \in \mathbb{R}^n, \ \alpha_k \in \mathbb{R} \tag{A8c}$$

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

$$\min_{z\in\mathbb{R}^n} f(z) = \sum_{t=0}^{N-1} \frac{1}{2}x_{t+1}^\top Q_{t+1}x_{t+1} + d_{xt+1}x_{t+1} + \frac{1}{2}u_t^\top R_t u_t + d_{ut}u_t \tag{A9a}$$

subject to

$$\begin{aligned}
x_{t+1} &= A_t x_t + B_t u_t, && t = 0,\dots,N-1 && \text{(A9b)}\\
x_0 &= \text{given} && && \text{(A9c)}\\
x^{\text{low}} &\le x_t \le x^{\text{high}}, && t = 1,\dots,N && \text{(A9d)}\\
u^{\text{low}} &\le u_t \le u^{\text{high}}, && t = 0,\dots,N-1 && \text{(A9e)}\\
-\Delta u^{\text{high}} &\le \Delta u_t \le \Delta u^{\text{high}}, && t = 0,\dots,N-1 && \text{(A9f)}\\
Q_t &\succeq 0 && t = 1,\dots,N && \text{(A9g)}\\
R_t &\succeq 0 && t = 0,\dots,N-1 && \text{(A9h)}
\end{aligned}$$

where

$$\begin{aligned}
u_t &\in \mathbb{R}^{n_u} && \text{(A9i)}\\
x_t &\in \mathbb{R}^{n_x} && \text{(A9j)}\\
\Delta u_t &= u_t - u_{t-1} && \text{(A9k)}\\
z^\top &= (x_1^\top,\dots,x_N^\top,u_0^\top,\dots,u_{N-1}^\top) && \text{(A9l)}
\end{aligned}$$

The subscript t denotes discrete time sampling instants.

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

$$\min_{z\in\mathbb{R}^n} f(z) = \sum_{t=0}^{N-1} \frac{1}{2}x_{t+1}^\top Q_{t+1}x_{t+1} + \frac{1}{2}u_t^\top R_t u_t \tag{A10a}$$

subject to

$$\begin{aligned}
x_{t+1} &= A_t x_t + B_t u_t && \text{(A10b)}\\
x_0 &= \text{given} && \text{(A10c)}
\end{aligned}$$

where

$$\begin{aligned}
u_t &\in \mathbb{R}^{n_u} && \text{(A10d)}\\
x_t &\in \mathbb{R}^{n_x} && \text{(A10e)}\\
z^\top &= (x_1^\top,\dots,x_N^\top,u_0^\top\dots,u_{N-1}^\top) && \text{(A10f)}
\end{aligned}$$

Theorem 2: *The solution of (A10) with $Q_t \succeq 0$ and $R_t \succ 0$ is given by*

$$u_t = -K_t x_t \tag{A11a}$$

where the feedback gain matrix is derived by

$$\begin{aligned}
K_t &= R_t^{-1}B_t^\top P_{t+1}(I + B_t R_t^{-1}B_t^\top P_{t+1})^{-1}A_t, && t = 0,\dots,N-1 && \text{(A11b)}\\
P_t &= Q_t + A_t^\top P_{t+1}(I + B_t R_t^{-1}B_t^\top P_{t+1})^{-1}A_t, && t = 0,\dots,N-1 && \text{(A11c)}\\
P_N &= Q_N && && \text{(A11d)}
\end{aligned}$$

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)