TTK4135 Optimization and Control
Solution to Final Exam - Spring 2012

Department of Engineering Cybernetics

1 Various Topics (36 %)

General Optimization Problems

We consider the general constrained optimization problem (A.1).

a (4 %) For the problem (A.1) to be convex,

- the objective function $f(x)$ must be convex,
- the equality constraint functions $c_i(x),\ i \in \mathcal{E}$ must be linear, and
- the inequality constraint functions $c_i(x),\ i \in \mathcal{I}$ must be concave.

b (2 %) Problem (A.1) can be formulated as the maximization problem

$$\max_{x\in\mathbb{R}^n} \quad -f(x) \tag{1a}$$
$$\text{s.t.} \quad c_i(x) = 0, \qquad i \in \mathcal{E} \tag{1b}$$
$$c_i(x) \geq 0, \qquad i \in \mathcal{I} \tag{1c}$$

c (4 %) When there are no inequality constraints in (A.1), i.e., $\mathcal{I} \in \emptyset$, a suitable merit function for use in an SQP algorithm is the ℓ_1 merit function

$$\phi_1(x;\mu) = f(x) + \mu\|c(x)\|_1 \tag{2}$$

with $\mu > 0$.

d (8 %) When (A.1) is unconstrained, i.e., $\mathcal{I} \in \emptyset$ and $\mathcal{E} \in \emptyset$, and f is convex, any local minimizer x^* is a global minimizer of f. In order to prove this, assume that x^* is a local but not a global minimizer. This means there exists a better minimizer $z \in \mathbb{R}^n$, such that $f(z) < f(x^*)$. We connect the two points x^* and z with a line segment:

$$x = \lambda z + (1-\lambda)x^*, \qquad \text{for some } \lambda \in (0,1] \tag{3}$$

Since f is convex, we have

$$f(x) \leq \lambda f(z) + (1-\lambda)x^* \leq f(x^*) \tag{4}$$

Any neighborhood $\mathcal{N}$ of x^* contains a piece of the line segment (3), so there will always be points $x \in \mathcal{N}$ at which (4) is satisfied. Hence, x^* is not a local minimizer.

Linear Programming

We consider the LP problem in the standard form (A.6).

e (4 %) A *basic feasible point* x for problem (A.6) is defined by the following:

- A subset $\mathcal{B} \subseteq \{1, \ldots, n\}$ can be defined as containing exactly m indices,
- $i \notin \mathcal{B} \Rightarrow x_i = 0$,
- the $m \times m$ matrix B defined by $B = [A_i]_{i \in \mathcal{B}}$, where A_i is the ith column of A, is nonsingular.

f (6 %) The LP problem

$$\begin{aligned} \min \quad & 3x_1 + 2x_2 + x_3 && (5a)\\ \text{s.t.} \quad & 2x_1 + 2x_2 + x_3 \leq 3 && (5b)\\ & x_1 - \ \ x_2 - x_3 \leq -1 && (5c)\\ & x \geq 0 && (5d) \end{aligned}$$

can be written as

$$\begin{aligned} \min \quad & 3x_1 + 2x_2 + x_3 && (6a)\\ \text{s.t.} \quad & 2x_1 + 2x_2 + x_3 + x_4 \qquad = 3 && (6b)\\ & x_1 - \ \ x_2 - x_3 \qquad + x_5 = -1 && (6c)\\ & x \geq 0 && (6d) \end{aligned}$$

which is in the standard form (A.6) with

$$\begin{aligned} c &= \begin{bmatrix} 3 & 2 & 1 & 0 & 0 \end{bmatrix}^\top, & x &= \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}^\top \\ A &= \begin{bmatrix} 2 & 2 & 1 & 1 & 0 \\ 1 & -1 & -1 & 0 & 1 \end{bmatrix}, & b &= \begin{bmatrix} 3 \\ -1 \end{bmatrix} \end{aligned} \tag{7}$$

g (8 %) The dual problem of the LP problem in standard form is given by

$$\begin{aligned} \max \quad & b^\top \lambda && (8a)\\ \text{s.t.} \quad & A^\top \lambda \leq c && (8b) \end{aligned}$$

In order to show that the KKT conditions for the dual problem are equal to the KKT conditions for the original LP problem, we first rewrite the dual problem as

$$\min -b^\top \lambda \qquad \text{s.t.} \qquad c - A^\top \lambda \geq 0 \tag{9}$$

and define the Lagringian for the problem as

$$\bar{\mathcal{L}}(\lambda, x) = -b^\top \lambda - x^\top (c - A^\top \lambda) \tag{10}$$

where $x \in \mathbb{R}^n$ are multipliers for the constraints. Differentiating with respect to λ and requiring the derivative to be zero gives

$$\nabla_\lambda \bar{\mathcal{L}} = -b + (x^\top A^\top)^\top = Ax - b = 0 \tag{11}$$

The KKT conditions (A.3) for the dual problem then becomes

$$Ax^* = b \tag{12a}$$
$$A^\top \lambda^* \leq c \tag{12b}$$
$$x^* \geq 0 \tag{12c}$$
$$x_i^*(c - A^\top \lambda^*)_i = 0, \quad i = 1, \dots, n \tag{12d}$$

The KKT conditions for (A.6) are

$$A^\top \lambda^* + s^* = c \tag{13a}$$
$$Ax^* = b \tag{13b}$$
$$x^* \geq 0 \tag{13c}$$
$$s^* \geq 0 \tag{13d}$$
$$s_i^* x_i^* = 0, \quad i = 1, \dots, n \tag{13e}$$

which are identical to (12) since $s = c - A^\top \lambda$.

2 Quadratic Programming (QP) (36 %)

We consider the QP problem in the standard form (A.7).

a (4 %) The active set $\mathcal{A}(x)$ for (A.7) is defined as

$$\mathcal{A}(x) = \{i \in \mathcal{E} \cup \mathcal{I} \mid a_i^\top x = b_i\} \tag{14}$$

which is the set of indices of the constraints for which equality holds at x.

b (8 %) When there are no inequality constraints in (A.7), the Lagrangean is

$$\mathcal{L}(x, \lambda) = f(x) - \lambda^\top (Ax - b) = \frac{1}{2} x^\top G x + x^\top c - \lambda^\top A x + \lambda^\top b \tag{15}$$

Requiring the gradient with respect to x of the Lagrangean to vanish at the solution x^* gives the condition

$$Gx^* + c - A^\top \lambda^* = 0 \tag{16}$$

This equation can be rearranged and combined with the condition

$$Ax^* = b \tag{17}$$

to form the KKT matrix equation

$$\begin{bmatrix} G & -A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix} \tag{18}$$

c (8 %) When there are no inequality constraints in (A.7), i.e., $\mathcal{I} \in \emptyset$, the necessary conditions for the vector x^* satisfying the KKT conditions to be the unique global solution of (A.7) are that A has full rank and that the reduced-Hessian matrix $Z^\top G Z$ is positive definite. Z is the $n \times (n - m)$ matrix whose columns are a basis for the nullspace of A.

d (6 %) When (A.7) is unconstrained, i.e., $\mathcal{I} \in \emptyset$ and $\mathcal{E} \in \emptyset$, and G is positive definite, the optimization problem is simply

$$\min f(x) = \frac{1}{2} x^\top G x + x^\top c, \quad G > 0 \tag{19}$$

- The Newton direction p_k^{N} is defined as

$$p_k^{\mathrm{N}} = (-\nabla^2 f_k)^{-1} \nabla f_k \tag{20}$$

  We have that

$$\nabla f_k = G x_k + c \tag{21}$$

  and that

$$\nabla^2 f_k = G \tag{22}$$

  Hence, the Newton direction for the problem is

$$p_k^{\mathrm{N}} = -G^{-1}(G x_k + c) = -x_k - G^{-1} c \tag{23}$$

- When G is positive definite, p_k^{N} is a descent direction since

$$(p_k^{\mathrm{N}})^\top \nabla f_k = (\nabla f_k)^\top (-\nabla^2 f_k)^{-\top} \nabla f_k = -(\nabla f_k)^\top G \nabla f_k < 0 \tag{24}$$

- When using the iteration algorithm $x_{k+1} = x_k + p_k^{\mathrm{N}}$, we have that

$$x_{k+1} = x_k + p_k^{\mathrm{N}} = x_k - x_k - G^{-1} c = -G^{-1} c \tag{25}$$

  which is the minimum since $G = G^\top > 0$.

e (10 %) A farmer wants to grow two different kinds of crop, C (e.g., carrot) and R (e.g., rutabaga) in a field F of size 100 000 m^2 which is available to him. Growing 1 tonne C in F requires an area of 4 000 m^2, whereas growing 1 tonne of R requires an area of 3 000 m^2. In addition, the two crops require different amounts of fertilizer. C requires 60 kg fertilizer per tonne grown, whereas R requires 80 kg fertilizer per tonne grown. The price of 1 kg fertilizer is 1. The market available for the farmer is very local and depends strongly upon the supply (measured in tonnes grown) of the crops. The selling price thus depends upon the supply.

- The price of crop C is $7000 - 200x_1$ per tonne, where x_1 is the number of tonnes grown of C.
- The price of crop R is $4000 - 140x_2$ per tonne, where x_2 is the number of tonnes grown of R.

Federal regulations allow the farmer to use at most a total of 2000 kg of fertilizer. The farmer wants to maximize the total profit.

The farmer's income from growing crop C is $(7000 - 200x_1)x_1$, where x_1 is the number of tonnes grown of C. His from growing crop R is $(4000 - 140x_2)x_2$, where x_2 is the number of tonnes grown of R. The only expense is fertilizer, which will cost him $60x_1 + 80x_2$. The total profit is the difference between income and expenses; the farmer wants to maximize

$$(7000 - 200x_1)x_1 + (4000 - 140x_2)x_2 - 60x_1 - 80x_2 \tag{26}$$

The size of his field leads to the constraint

$$4000x_1 + 3000x_2 \leq 100000 \tag{27}$$

while the federal regulations on fertilizer use leads to the constraint

$$60x_1 + 80x_1 \leq 2000 \tag{28}$$

Since production must be nonnegative,

$$x \geq 0 \tag{29}$$

Hence, we have the quadratic programming problem

$$\begin{aligned}
\min_{x} \quad & f(x) = (7000 - 200x_1)x_1 + (4000 - 140x_2)x_2 - 60x_1 - 80x_2 && \text{(30a)}\\
\text{s.t.} \quad & 4000x_1 + 3000x_2 \leq 100000 && \text{(30b)}\\
& 60x_1 + 80x_1 \leq 2000 && \text{(30c)}\\
& x \geq 0 && \text{(30d)}
\end{aligned}$$

3 Optimal Control and MPC (28 %)

We consider the optimal control problem for a linear dynamic system described by (A.9) on a finite horizon of length n. Q_i, $1 \leq i \leq n$, and S are symmetric positive semidefinite matrices ($Q_i \geq 0$, $S \geq 0$), whereas P_i, $1 \leq i \leq n$, are symmetric positive definite matrices ($P_i > 0$).

a (10 %) We assume that there are no inequality constraints in this problem, i.e., we remove (A.9e) and (A.9f).

- The Lagrangean function for (A.9a)–(A.9d) can be written

$$\mathcal{L} = f_0 - \sum_{i=0}^{n-1} \lambda_{i+1}(x_{i+1} - A_i x_i - B_i u_i) \tag{31}$$

- The KKT conditions for (A.9a)–(A.9d) can be written

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial u_i} &= P_i(u_i - u_{\text{ref},i}) + B_i^\top \lambda_{i+1} = 0, && 0 \leq i \leq n-1 && \text{(32a)}\\
\frac{\partial \mathcal{L}}{\partial x_i} &= Q_i(x_i - x_{\text{ref},i}) - \lambda_i + A_i^\top \lambda_{i+1} = 0, && 0 \leq i \leq n-1 && \text{(32b)}\\
\frac{\partial \mathcal{L}}{\partial x_n} &= S(x_n - x_{\text{ref},n}) - \lambda_n = 0 && && \text{(32c)}\\
x_{i+1} &= A_i x_i + B_i u_i, && 0 \leq i \leq n-1 && \text{(32d)}
\end{aligned}$$

- The resulting controller $u_i = K_i x_i$ is a linear, time-varying, state-feedback controller.

b (6 %) We now consider the infinite horizon (the limit as n tends to infinity) optimal control problem

$$\min \quad f^\infty = \frac{1}{2}\sum_{i=0}^{\infty}\left\{x_i^\top Q x_i + u_i^\top P u_i\right\} \tag{33a}$$

$$\text{s.t.} \quad x_{i+1} = Ax_i + Bu_i, \quad 0 \le i \le \infty \tag{33b}$$

- The controller $u_i = Kx_i$ is computed by solving

$$K = -P^{-1}B^\top R(I + BP^{-1}B^\top R)^{-1}A \tag{34a}$$

$$R = Q + A^\top R(I + BP^{-1}B^\top R)^{-1}A \tag{34b}$$

$$R = R^\top \ge 0 \tag{34c}$$

- The conditions on A, B, and D ($Q = D^\top D$) that satisfied for asymptotic stability of optimal closed-loop system are
  - (A, B) stabilizable; (A, B) is stabilizable if all the uncontrollable modes are asymptotically stable.
  - (A, D) detectable; (A, C) is detectable if all the unobservable modes are asymptotically stable.

The remainder of this problem concerns MPC where the constraints (A.9e) and (A.9f) must be taken into account.

c (4 %) Model predictive control is a form of control in which the current control action is obtained by solving, at each sampling instant, a finite horizon open-loop optimal control problem, using the current state of the plant as the initial state; the optimization yields an optimal control sequence and the first control in this sequence is applied to the plant.[1]

d (4 %)

- Some reasons for the industrial success of MPC are (1) inherent multivariable control, (2) handling of constraints, both on inputs and states, (3) the possibility of operation closer to constraints, usually leading to a more profitable process, and (4) understandable and intuitive theory.
- Reasons for why the majority of MPC applications us based on linear models, despite all practical systems exhibiting nonlinear behavior include (1) Most plants operate close to some operating point and the nonlinear behavior such points can be approximated by linear models, and (2) Solving nonlinear optimization problems online requires has a high computational cost and may take a long time. Additionally, it may be difficult to guarantee that a solution to the optimization problem is found within reasonable time, if it is possible to guarantee that one will be found at all. are required

[1] Mayne, D. Q., Rawlings, J. B., Rao, C. V. and Scokaert, P. O. M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814.

e (4 %) In order to ensure that a feasible solution exists, we can add slack variables $0 \leq \epsilon_i \in \mathbb{R}^l$ to the state (or output) inequality constraints:

$$X_L - \epsilon_i \leq x_i \leq X_U + \epsilon_i \tag{35}$$

A term $\rho^\top \epsilon_k$, $0 \leq \rho \in \mathbb{R}^l$ would have to be added to the objective function. This is an exact penalty method, provided ρ is large enough. (A quadratic penalty on ϵ_i is also possible.)

Appendix

Part 1 Optimization Problems and Optimality Conditions

A general formulation for constrained optimization problems is

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

where f and 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:

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

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:

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

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

QP problem in standard form:

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

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:

$$x_{k+1} = x_k + \alpha_k p_k \tag{A.8a}$$
$$x_0 \text{ given} \tag{A.8b}$$
$$x_k, p_k \in \mathbb{R}^n, \;\; \alpha_k \in \mathbb{R} \tag{A.8c}$$

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

$$\begin{aligned}\min\ f_0 = \frac{1}{2}\sum_{i=0}^{n-1} &\big\{(y_i - y_{\mathrm{ref},i})^\top Q_i (y_i - y_{\mathrm{ref},i}) \\ &+ (u_i - u_{i-1})^\top P_i (u_i - u_{i-1})\big\} \\ &+ \frac{1}{2}(y_n - y_{\mathrm{ref},n})^\top S (y_n - y_{\mathrm{ref},n})\end{aligned} \tag{A.9a}$$

subject to equality and inequality constraints

$$x_{i+1} = A_i x_i + B_i u_i,\ 0 \le i \le n-1 \tag{A.9b}$$
$$y_i = H x_i \tag{A.9c}$$
$$x_0 = \text{given (fixed)} \tag{A.9d}$$
$$U_L \le u_i \le U_U,\ 0 \le i \le n-1 \tag{A.9e}$$
$$Y_L \le y_i \le Y_U,\ 1 \le i \le n \tag{A.9f}$$

where system dimensions are given by

$$u_i \in \mathbb{R}^m \tag{A.9g}$$
$$x_i \in \mathbb{R}^l \tag{A.9h}$$
$$y_i \in \mathbb{R}^j \tag{A.9i}$$

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 2: *Assume that* $x_{\mathrm{ref},i} = 0$, $u_{\mathrm{ref},i} = 0$, $0 \le i \le n$ *and that* $H = I$, *i.e.,* $y_i = x_i$. *The solution of* (A.9a), (A.9b), *and* (A.9d) *is given by* $u_i = K_i x_i$, $0 \le i \le n-1$, *where the feedback gain matrix is derived by*

$$K_i = -P_i^{-1} B_i^T R_{i+1} (I + B_i P_i^{-1} B_i^T R_{i+1})^{-1} A_i,\ 0 \le i \le n-1 \tag{A.10a}$$
$$R_i = Q_i + A_i^T R_{i+1} (I + B_i P_i^{-1} B_i^T R_{i+1})^{-1} A_i,\ 0 \le i \le n-1 \tag{A.10b}$$
$$R_n = S \tag{A.10c}$$