TTK4135 Optimization and Control
Solution to Final Exam — Spring 2016

Department of Engineering Cybernetics

1 Linear Programming (40 %)

Problem formulation

a (12 %) The problem can be formulated as follows:

$$\begin{aligned} &\min -7000x_1 - 6000x_2, \\ \text{subject to} \quad & 4000x_1 + 3000x_2 \leq 100000, \\ & 60x_1 + 80x_2 \leq 2000, \\ & x_1, x_2 \geq 0. \end{aligned}$$

By introducing slack variables, we have:

$$\begin{aligned} &\min -7000x_1 - 6000x_2, \\ \text{subject to} \quad & 4000x_1 + 3000x_2 + x_3 = 100000, \\ & 60x_1 + 80x_2 + x_4 = 2000., \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

Therefore, the matrices and vectors are as follows:

$$\begin{aligned} c^T &= \begin{bmatrix} -7000 & -6000 & 0 & 0 \end{bmatrix}, \\ A &= \begin{bmatrix} 4000 & 3000 & 1 & 0 \\ 60 & 80 & 0 & 1 \end{bmatrix}, \\ b &= \begin{bmatrix} 100000 \\ 2000 \end{bmatrix}. \end{aligned}$$

Simplex method

b (4 %)
- No, The simplex method doesn not use gradients,
- Yes.

Convexity

c (3 %)
- Yes, it is a convex problem,
- Linear objective function is a convex function,
- Linear constraints which gives a feasible convex set.

LU factorization

d (7 %)
- $C = LU$ where L is unit lower triangular and U is upper triangular,
- We have $Cz = d$. Therefore $LUz = d$. Set $Uz = y$.
  First, solve $Ly = d$ by performing triangular forward-substitution to obtain y.
  Second, solve $Uz = y$ by performing triangular back-substitution to obtain z.

QP problem

e (8 %)

- The objective function can be formulated by:

$$\begin{aligned} f(\mathbf{x}) &= -(7000 - 200x_1)x_1 - (6000 - 140x_2)x_2, \\ &= 200x_1^2 - 7000x_1 + 140x_2^2 - 6000x_2, \\ &= \frac{1}{2}\mathbf{x}^T G\mathbf{x} + c^T\mathbf{x},\ G = \begin{bmatrix} 400 & 0 \\ 0 & 280 \end{bmatrix}, c = \begin{bmatrix} -7000 & -6000 \end{bmatrix}^T. \end{aligned}$$

  Therefore, the problem can be formulated as follows:

$$\begin{aligned} & \min f(\mathbf{x}) \\ \text{subject to} \quad & 4000x_1 + 3000x_2 \leq 100000, \\ & 60x_1 + 80x_2 \leq 2000, \\ & x_1, x_2 \geq 0. \end{aligned}$$

- QP problem
- Yes, it is a convex problem, since $G \succ 0$ and the feasible set is convex.

Sensitivity

f (6 %) We need to use $\lambda_2^* = 21.4$ which gives the sensitivity of the corresponding constraint w.r.t. profit. Therefore, the new profit will be: $185714 + 21.4 = 185735$.

2 MPC and optimal control (40 %)

a (4 %)

- Advantages: Smaller problems and less complex formulations.
- Disadvantages: It cannot handle a system with significant dynamics.

b (26 %)

b1 Open loop because it does not include feedback control.

b2 $\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t$.

b3 This is a limit on change of control input and/or limit wear and tear of control input.

b4 QP.

b5 NLP problem, and a suitable solution algorithm is SQP.

b6 We normally have tighter bounds on the state variables than reality since some of the constraints are hard and must always be satisfied. Therefore, the state constraints (A.9d) may be violated for all the time (e.g., if a severe disturbance occurs between time step $t-1$ and t and moves the state beyond the state constraint limit, then no feasible point exists at t). In this case a feasible point may not exist and a control input may not be available. This is an unacceptable situation. To avoid this, we soften the state constraint (A.9d) by using slack variables as below:

$$\begin{aligned}\min_{z\in\mathbb{R}^n} f(z) = \sum_{t=0}^{N-1} & \frac{1}{2}x_{t+1}^{\top}Q_{t+1}x_{t+1} + d_{xt+1}x_{t+1} \\ & + \frac{1}{2}u_t^{\top}R_tu_t + d_{ut}u_t + \rho^T\epsilon + \frac{1}{2}\epsilon^T S\epsilon\end{aligned} \tag{1a}$$

subject to

$$x_{t+1} = A_tx_t + B_tu_t, \qquad t = 0,\dots,N-1 \tag{1b}$$

$$x_0 = \text{given} \tag{1c}$$

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

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

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

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

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

where

$$\epsilon \in \mathbb{R}^{n_x} \ge 0 \tag{1i}$$

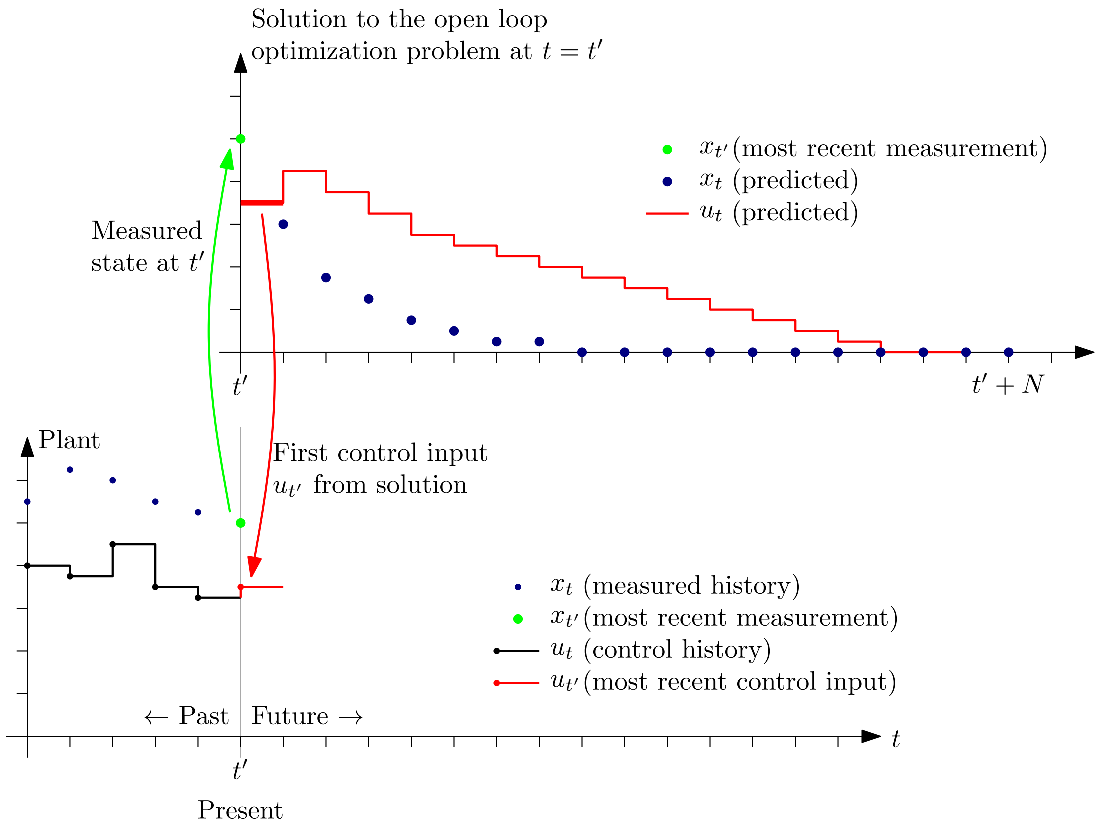
$$\rho \in \mathbb{R}^{n_x} \ge 0 \tag{1j}$$

$$S = \text{diag}\{s_1, \cdots, s_{n_x}\},\ s_i \ge 0,\ i = \{1, 2, \cdots, n_x\}. \tag{1k}$$

Two positive terms $\rho^T\epsilon$ and/or $\frac{1}{2}\epsilon^T S\epsilon$ are added to the original QP problem. These are both positive terms, hence there is a desire to derive these terms to zero. More precisely, the slack variables should be nonzero if the corresponding constraints are violated. Adding $\rho^T\epsilon$ is like adding a penalty function and if ρ is chosen big enough, then the solution is exact.

b7 Number of optimization variables: $(5+2) \times 12 = 84$.
Reduced space formulation, removes all states variables by using (1b). Therefore, the number of optimization variables is reduced to $2 \times 12 = 24$.
Full space: many variables, but often a lot of sparsity in matrices. This can be explored by a solver.
Reduced space: less variables, but normally dense matrices.

c (10 %) 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.Figure. 1 illustrates the principle of MPC.



Figur 1: MPC principle.

3 Linear Programming (40 %)

Problem formulation

a (8 %)

$$
\begin{aligned}
f(x_1, x_2) &= (a - x_1)^2 + b(x_2 - x_1^2)^2, \\
\nabla f(x_1, x_2) &= \begin{bmatrix} -2(a - x_1) + 2b(x_2 - x_1^2)(-2x_1) \\ 2b(x_2 - x_1^2) \end{bmatrix}, \\
\nabla^2 f(x_1, x_2) &= \begin{bmatrix} 2 - 4b(x_2 - x_1^2) + 8bx_1^2 & -4bx_1 \\ -4bx_1 & 2b \end{bmatrix} \\
&= \begin{bmatrix} 2 - 4bx_2 + 12bx_1^2 & -4bx_1 \\ -4bx_1 & 2b \end{bmatrix}.
\end{aligned}
$$

b (6 %) $a = 1,\ b = 2, \mathbf{x}^* = (1, 1)^T$

$$
\begin{aligned}
\nabla f(\mathbf{x}^*) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\
\nabla^2 f(\mathbf{x}^*) &= \begin{bmatrix} 17 & -8 \\ -8 & 4 \end{bmatrix},
\end{aligned}
$$

The Hessian matrix is positive deffinite, as $(det(\nabla^2 f(\mathbf{x}^*)) > 0)$. Therefore, both conditions are satisfied.

c (6 %)

$$
\min (a - x_1)^2 + b(x_2 - x_1^2)^2, \text{subject to} x_1 \geq 0,\ x_2 \geq 0.
$$

- Feasible set is convex,
- Optimization problem is non-convex since Rosenbrock function is a non-convex function.

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:

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

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

QP problem in standard form:

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

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 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

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

where

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

The subscript t denotes discrete time sampling instants.

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

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

subject to

$$\begin{aligned}
x_{t+1} &= A_t x_t + B_t u_t & \text{(A.10b)}\\
x_0 &= \text{given} & \text{(A.10c)}
\end{aligned}$$

where

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

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

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

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

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

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)