Delft University of Technology

Networked and Distributed Control Systems

SC42100

Assignment 2

Authors:
Winnifred Noorlander(4307925)
Django Beek (4281918)

April 28, 2020

Problem 1

(a) The following convex optimization problem:

$$\begin{aligned} \min_{\theta_1,\theta_2} \quad & f_1(\theta_1) + f_2(\theta_2) \\ \text{s.t.} \quad & \theta_1 \in \Theta_1, \quad \theta_2 \in \Theta_2 \\ & h_1(\theta_1) + h_2(\theta_2) \leq 0 \end{aligned}$$

with Θ_1, Θ_2 representing convex sets, can be decomposed using Primal Decomposition. The objective function already has a separable structure and can be trivially decomposed into two objective functions. For the complicating constraint however, an extra fixed variable, c, is introduced to aid the decoupling as follows [Boyd et al., 2015]. Rewriting the constraint to $h_1 \leq -h_2$, and defining c to be

$$h_1 \leq c, \tag{1}$$

then $h_1 \leq c \leq -h_2$ and thus

$$h_2 \leq -c. \tag{2}$$

The resulting two subproblems are:

$$\begin{aligned} \min_{\theta_1} \quad & f_1(\theta_1) & \qquad \min_{\theta_2} \quad & f_2(\theta_2) \\ \text{s.t.} \quad & \theta_1 \in \Theta_1 & \qquad \text{s.t.} \quad & \theta_2 \in \Theta_2 \\ & h_1(\theta_1) \leq c & & -h_2(\theta_2) \leq c \end{aligned} \tag{3}$$

These hierarchically lower level problems fall under the master problem, which optimizes the newly introduced coupling variable c as follows:

$$\min_{c} \quad \phi_1(c) + \phi_2(c) \tag{4}$$

The original problem is convex, thus so is the master problem, which includes solving the two sub problems each iteration in parallel.

(b) To show that $-\lambda^*$ is a subgradient of p at z, another point $\tilde{z}$ can be introduced to evaluate p with, in the hope that this holds for all other points $\tilde{z}$ in domain p. The Lagrangian becomes:

$$L(\theta, \lambda) = f(\theta) + \lambda(h(\theta) - \tilde{z}). \tag{5}$$

Then considering p at $\tilde{z}$ by following [Boyd et al., 2015],

$$\begin{aligned} p(\tilde{z}) &= \sup_{\lambda \succeq 0} \inf_{x \in X} (f(x) + \lambda^T(h(x) - \tilde{z})) \\ &\geq \inf_{x \in X} (f(x) + \lambda^T(z)(h(x) - \tilde{z})) \\ &= \inf_{x \in X} (f(x) + \lambda^T(z)(h(x) - z + z - \tilde{z})) \\ &= \inf_{x \in X} (f(x) + \lambda^T(z)(h(x) - z)) + \lambda^T(z)(z - \tilde{z})) \\ &= \phi(z) + (-\lambda(z))^T(\tilde{z} - z). \end{aligned}$$

In which $\lambda = \lambda^*$ for abbreviation. This results in:

$$p(\tilde{z}) \geq \phi(z) + (-\lambda(z))^T(\tilde{z} - z),$$

which holds for all points $\tilde{z}$ in domain p, so $-\lambda(z)$ is a subgradient of p.

Problem 2

Firstly, since matrix W is doubly stochastic, balanced, and strongly connected, the proposed limit of $\varphi \to \infty$ exists. The doubly stochastic converges to a consensus (all elements in W converge to $\frac{1}{N}$) with this limit, over all agents because it is the strongly connected. Then, following [Johansson et al., 2008], for l $i = 1, ..., N$,

$$\lim_{\varphi \to \infty} \sum_{j=1}^{N} \Big([W^{\varphi}]_{ij} (\theta_0^j - \alpha_0 g^j(\theta_0^j) \Big) = \frac{1}{N} \sum_{j=1}^{N} \Big((\theta_0^j - \alpha_0 g^j(\theta_0^j) \Big). \tag{6}$$

With the initial state of the projected consensus denoted as follows:

$$\bar{\theta}_1 = \mathcal{P}_{\Theta} \left[\frac{1}{N} \sum_{j=1}^{N} \Big(\theta_0^j - \alpha_0 g^j(\theta_0^j) \Big) \right], \tag{7}$$

and as earlier stated that at every iteration all agents will converge to a consensus, thus posses the same value $\bar{\theta}_1$ and the local subgradients g^j will be evaluated at the same point. Translating Equation (7) into an updatable structure gives:

$$\begin{aligned} \bar{\theta}_{k+1} &= \mathcal{P}_{\Theta} \left[\frac{1}{N} \sum_{j=1}^{N} \Big(\theta_k^j - \alpha_k g^j(\theta_k^j) \Big) \right] \\ &= \mathcal{P}_{\Theta} \left[\Big(\theta_k^j - \alpha_k \frac{1}{N} \sum_{j=1}^{N} g^j(\theta_k^j) \Big) \right]. \end{aligned} \tag{8}$$

Problem 3

The cost function of the to be optimization problem that must be minimized is defined as:

$$\begin{aligned} J(x_{i,0}, u_i) = \sum_t & x_1(t)^T x_1(t) + u_1(t)^T u_1(t) + x_2(t)^T x_2(t) + u_2(t)^T u_2(t) \\ & + x_3(t)^T x_3(t) + u_3(t)^T u_3(t) + x_4(t)^T x_4(t) + u_4(t)^T u_4(t) \\ & = ||F_i x_{i,0} + h_i u_i||^2 + ||u_i||^2 \\ & = u_i^T (h_1^T h_1 + I_{10}) u_1 + 2 x_{i,0}^T F_i x_i^T h_i u_i + x_{i,0}^T F_i^T F_i x_{i,0} \end{aligned} \tag{9}$$

with:

$$F = \begin{bmatrix} I \\ A_i \\ A_i^2 \\ A_i^3 \\ A_i^4 \end{bmatrix}, H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ B_i & 0 & 0 & 0 & 0 \\ A_i B_i & B_i & 0 & 0 & 0 \\ A_i^2 B_i & A_i B_i & B_i & 0 & 0 \\ A_i^3 B_i & A_i^2 B_i & A_i B_i & B_i & 0 \\ A_i^4 B_i & A_i^3 B_i & A_i^2 B_i & A_i B_i & B_i \end{bmatrix}, u_i = \begin{bmatrix} u_i(0) \\ u_i(1) \\ u_i(2) \\ u_i(3) \end{bmatrix} \tag{10}$$

subject to the constraints:

$$\begin{aligned} |u_i(t)| &\leq \frac{u_{max}}{T_{final}}, \forall i, t \\ x_i(Tfinal) &= x_f, \forall i \end{aligned} \tag{11}$$

with $u_{max} = 100$ and $T_{Final} = 5$. Next to this the second linear constraint is rewritten into three separate constraints:

$$\begin{aligned} x_1(Tfinal) &= x_2(Tfinal) \\ x_2(Tfinal) &= x_3(Tfinal) \\ x_3(Tfinal) &= x_4(Tfinal) \end{aligned} \tag{12}$$

This physically means all planes must be in the same position at T_{final}. Now the constraints are now, the next step is to set up the Lagrangian of the problem:

$$\begin{aligned} L(u_1, u_2, u_3, u_4, \lambda) = &\sum_i \sum_t (J_1(x_0, u_1) + J_2(x_0, u_2) + J_3(x_0, u_3) + J_4(x_0, u_4) \\ &+ \lambda_1(x_1(Tfinal) - x_2(Tfinal)) + \lambda_2(x_2(Tfinal) - x_3(Tfinal)) \\ &+ \lambda_3(x_3(Tfinal) - x_4(Tfinal)) \end{aligned} \tag{13}$$

Wherein first the four subproblems are updated with constant λ's after which the λ's are updated using the sub-gradient from the subproblems. The update rule is given below:

$$\lambda_{1,k+1} = \lambda_k + \alpha(x_1(Tfinal) - x_2(Tfinal)) \tag{14}$$

The other λ 's are update in the same way. The initial values for λ are set to zero.
Using Quadratic Programming (`quadprog` command in matlab) and only part of the objective function per plane with the linear constraints as in Equation 11:

$$\min_u = 0.5 * u^T H u + h1u \tag{15}$$

Setting the step size to a relatively small number of $\alpha = 1$, the two problems are iteratively solved. Convergence of the states can be seen in Figure 1

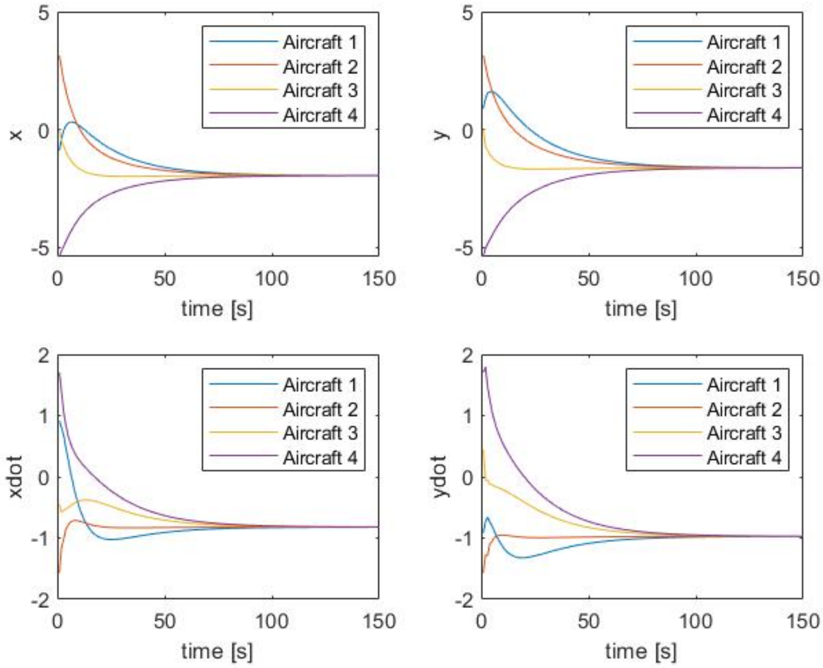

Figure 1: Convergence of all four planes on all four states

Problem 4

a.

From the relation below the next iteration is calculated using the initial point (u_1^p, u_2^p)

$$V(u_1, u_2) = \frac{1}{2}u_1^T H_{11} u_1 + u_2^T H_{21} u_1 + u_1^T H_{12} u_2 + u_2^T H_{22} u_2 + c_1^T u_1 + c_2^T u_2 + d \quad (16)$$

The derivative of the cost function with respect to u_1 is given by

$$\begin{aligned} \frac{d(V(u_1, u_2))}{d(u_1)} &= \frac{1}{2}(2u_1^T H_{11} + u_2^T H_{21} + u_2^T H_{12}^T) + c_1^T \\ \frac{d(V(u_1, u_2))}{d(u_2)} &= \frac{1}{2}(2u_2^T H_{22} + u_1^T H_{12} + u_1^T H_{21}^T) + c_2^T \end{aligned} \quad (17)$$

Given that the matrix H is positive definite, it follows H is symmetric and the entries $H_{12} = H_{21}^T$. Simplifying the derivatives to

$$\begin{aligned} \frac{d(V(u_1, u_2))}{d(u_1)} &= u_1^T H_{11} + u_2^T H_{12} + c_1^T \\ \frac{d(V(u_1, u_2))}{d(u_2)} &= u_2^T H_{22} + u_2^T H_{21} + c_2^T \end{aligned} \quad (18)$$

Setting the equations to zero and solving for u_1^p and for u_2^p gives the optimal solutions at the next time instances u_1^{p+1} and u_2^{p+1}.The relations are smoothend by making use of the fact H is symmetric, it then holds that $H^{-1} = (H^{-1})^T$.

$$
\begin{aligned}
u_1^{p+1} &= -(H_{11}^{-1})^T(H_{12}^T u_2^p + c_1) \\
u_2^{p+1} &= -(H_{22}^{-1})^T(H_{21}^T u_1^p + c_2)
\end{aligned}
\tag{19}
$$

b.

In this exercise it is shown that the iteration matrix A

$$
A = \begin{bmatrix} 0 & -H_{11}^{-1}H_{12} \\ -H_{22}^{-1}H_{21} & 0 \end{bmatrix}
\tag{20}
$$

is stable and hence the optimization procedure converges. To do so it must be noticed that the determinant of H is larger than zero because of the positive definite property.

$$
\begin{aligned}
det(H) &= H_{11}H_{22} - H_{12}H_{21} > 0 \\
H_{11}H_{22} &> H_{12}H_{21}
\end{aligned}
\tag{21}
$$

The eigenvalues of A, λ, are given by

$$
\begin{aligned}
0 &= \lambda^2 - H_{11}^{-1}H_{12}H_{22}^{-1}H_{21} \\
\lambda &= \sqrt{\frac{H_{12}H_{21}}{H_{11}H_{22}}}
\end{aligned}
\tag{22}
$$

Hence it is proven that

$$
|eig(A)| < 1
\tag{23}
$$

c.

Given that the eigenvalues of A are smaller than one and by using the geometric series of matrices given as

$$
\begin{aligned}
lim_{p\to\infty} u^{p+1} &= A^{\infty}u^0 + \sum_{p=0}^{\infty}(A^p)b \\
&= (I - A^{-1})b
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
I - A^{-1} &= (I - A)^{-1} \\
&= \begin{bmatrix} I & H_{11}^{-1}H_{12} \\ H_{22}^{-1}H_{21} & I \end{bmatrix}^{-1}
\end{aligned}
\tag{25}
$$

with the constant b given as

$$
\begin{aligned}
b &= \begin{bmatrix} -H_{11}^{-1}c_1 \\ -H_{22}^{-1}c_2 \end{bmatrix} \\
&= \begin{bmatrix} -H_{11}^{-1} & 0 \\ 0 & -H_{22}^{-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}
\end{aligned}
\tag{26}
$$

Now rewriting using the Schur complement over entry H_{22} and the matrix inversion lemma.

$$
\begin{aligned}
(I-A^{-1})b &= \begin{bmatrix} (H_{11}-H_{12}H_{22}^{-1}H_{21})^{-1} & -(H_{11}-H_{12}H_{22}^{-1}H_{21})^{-1}H_{21}H_{22}^{-1} \\ -H_{22}^{-1}H_{21}(H_{11}-H_{12}H_{22}^{-1}H_{21})^{-1} & H_{22}^{-1}+H_{22}^{-1}H_{21}(H_{11}-H_{12}H_{22}^{-1}H_{21})^{-1}H_{12}H_{22}^{-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
&= \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}
\end{aligned}
\tag{27}
$$

that the solution is

$$u^* = -H^{-1}c \tag{28}$$

when the iteration converges.

Problem 5

a.

To show that the cost function is monotonically decreasing when optimizing one variable at the time, one can look at the derivative of the difference in step size.

First of all let's take into account that:

$$u^{p+1} = Au^p + b \tag{29}$$

Giving V at the different points as:

$$
\begin{aligned}
V(u^p) &= \frac{1}{2}u^{pT}Hu^p + c^Tu^p + d \\
V(u^{p+1}) &= \frac{1}{2}u^{p+1T}Hu^{p+1} + c^Tu^{p+1} + d \\
&= \frac{1}{2}(u^{pT}A^THAbu^p + b^THb + u^{pT}A^THb + b^THAu^p) + c^Tb + c^TAu^p + d
\end{aligned}
\tag{30}
$$

The difference between these is given as

$$V(u^{p+1}) - V(u^p) = \frac{1}{2}(u^{pT}A^THAu^p + b^THb + u^{pT}A^THb + b^THAu^p) + c^Tb + c^TAu^p - c^Tu^p \tag{31}$$

with A^T given as

$$A^T = \begin{bmatrix} 0 & -H_{21}H_{22}^{-1} \\ -H_{12}H_{11}^{-1} & 0 \end{bmatrix} \tag{32}$$

Looking at the shape of the derivative to u of this cost function tells if it is monotonically decreasing. The derivative to u is given by

$$
\begin{aligned}
\frac{d(V(u^{p+1})-V(u^p))}{d(u^p)} &= u^{pT}A^THA + b^THA - u^{pT}H + c^T(A-I) \\
&= u^{pT}(\begin{bmatrix} 0-H_{21}H_{22}^{-1} & \\ -H_{12}H_{11}^{-1} & 0 \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} 0 & -H_{11}^{-1}H_{12} \\ 0 & -H_{22}^{-1}H_{21} \end{bmatrix} - \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}) \\
&+ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \begin{bmatrix} -H_{11}^{-1} & 0 \\ 0 & -H_{22}^{-1} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} 0 & -H_{11}^{-1}H_{12} \\ 0 & -H_{22}^{-1}H_{21} \end{bmatrix} \\
&= u^{pT}(\begin{bmatrix} H_{11} & 0 \\ 0 & H_{22} \end{bmatrix} \\
&\begin{bmatrix} H_{11}^{-1}H_{12}H_{22}^{-1}H_{21} - I & H_{11}^{-1}H_{12}H_{22}^{-1}H^{21H_{11}^{-1}H_{12}-H_{11}^{-1}H_{12}} \\ H_{22}^{-1}H_{21}H_{11}^{-1}H_{12}H_{22}^{-1}H_{21} - H_{22}^{-1}H_{21} & H_{22}^{-1}H_{21}H_{11}^{-1}H_{12} - I \end{bmatrix} \\
&+ c^T(\begin{bmatrix} -H_{11}^{-1} & 0 \\ 0 & -H_{22}^{-1} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} 0 & -H_{11}^{-1}H_{12} \\ -H_{22}^{-1}H_{21} & 0 \end{bmatrix} \\
&+ \begin{bmatrix} I & -H_{11}^{-1}H_{12} \\ -H_{22}^{-1}H_{21} & I \end{bmatrix}) \\
&= u^{pT}(A + A^2 + A^3 - I) + c^T(\begin{bmatrix} I & -H_{11}^{-1}H_{12} \\ -H_{22}^{-1}H_{21} & I \end{bmatrix} \begin{bmatrix} 0 & -H_{11}^{-1}H_{21} \\ -H_{22}^{-1}H_{21} & 0 \end{bmatrix} \\
&+ \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \\
&= u^{pT}H(A^2 - I) + c^T(A^2 - I)
\end{aligned} \tag{33}
$$

If the iteration converges, the step size becomes zero, meaning $u^{p+1} = u^p = u^*$. Furthermore it is known that —eig(A)— <1, so $|eig(A^2| < 1$, so $|eig(A-I)| < 0$. It can than be concluded that A is negative definite. Now substituting

$$
\begin{aligned}
u* &= -H^-1c \\
c &= -Hu*
\end{aligned} \tag{34}
$$

into 33 gives the following result:

$$
\begin{aligned}
\frac{d(V(u^{p+1})-V(u^p))}{d(u^p)} &= u^{pT}H(A^2 - I) - u^{*T}H(A^2 - I) \\
&= (u^{pT} - u^{*T})H(A^2 - I)
\end{aligned} \tag{35}
$$

Furthermore $u^p - u^* < 0$ gives a positive derivative and $u^p - u^* > 0$ gives a negative derivative. The derivative therefore has the shape of a downward opening parabola. Combining this with the property that H is positive definite and thus that $H(A^2-I) < 0$ it can finally be said that the derivative is monotonically decreasing when only one variable is optimized and $u^p \neq u^*$.

b.

To define the size of the decrease, first the coordination system is recalibrated by setting u^* as the centre of the coordinate system. This means replacing u by $u^\bullet = u - u^*$.

Substituting this we get $c = 0$ and hence $b = 0$. The step size function is then simplified to

$$V(u^{\bullet(p+1)} - V(u^{\bullet(p)} = 0.5 u^{\bullet pT}(A^T H A - H) u^{\bullet p} \tag{36}$$

The given definition for P can be fully written out and one then finds:

$$\begin{aligned} P &= H D^{-1} \check{H} D^{-1} H \\ &= -(A^T H A - H) \end{aligned} \tag{37}$$

The step sizes then becomes:

$$V(u^{\bullet(p+1)} - V(u^{\bullet(p)} = -0.5 u^{\bullet pT} P u^{\bullet p} \tag{38}$$

Switching back to the original coordinate system gives the desired result.

References

[Boyd et al., 2015] Boyd, S., Xiao, L., Mutapcic, A., and Mattingley, J. (May 2015). Snotes on decomposition methods. Notes for EE364B, Stanford University, Winter 2006-07.

[Johansson et al., 2008] Johansson, B., Keviczky, T., Johansson, M., and Johansson, K. H. (2008). Subgradient methods and consensus algorithms for solving convex optimization problems. *47th IEEE Conference on Decision and Control*.