Delft University of Technology

# Networked and Distributed Control
# Assignment 2


*Author:*

Joost Jeschke (4309111) and Miranda van Duijn (4355776)


June 26, 2019

# Problem 1

a) The following convex optimization problem with a complicating constraint is considered:

$$\begin{aligned} \underset{\theta_1,\theta_2}{\text{minimize}} \quad & f_1(\theta_1) + f_2(\theta_2) \\ \text{subject to} \quad & \theta_1 \in \Theta_1, \quad \theta_2 \in \Theta_2 \\ & h_1(\theta_1) + h_2(\theta_2) \leq 0 \end{aligned}$$

The problem is coupled due to the second constraint, the complicating constraint. By introducing the real coupling variable $t$, which represents the amount of the resources allocated to the first subproblem [1]. $-t$ is allocated to the second subproblem. Using the primal decomposition, the subproblems are:

$$\begin{aligned} \underset{\theta_1}{\text{minimize}} \quad & f_1(\theta_1) \\ \text{subject to} \quad & \theta_1 \in \Theta_1 \\ & h_1(\theta_1) \leq t \end{aligned}$$

$$\begin{aligned} \underset{\theta_2}{\text{minimize}} \quad & f_2(\theta_2) \\ \text{subject to} \quad & \theta_2 \in \Theta_2 \\ & h_2(\theta_2) \leq -t \end{aligned}$$

The subproblems can be solved for a fixed value of $t$. Let $\phi_1(t)$ and $\phi_2(t)$ denote the optimal solutions to the two subproblems. Then the unconstrained master problem becomes:

$$\min_t \phi_1(t) + \phi_2(t)$$

The role of the master problem is to set the pricing strategy, meaning that the master problem sets the price of the resources of each subproblem.

b) It needs to be proven that for any other point $\tilde{z} \in \mathbb{R}^M$, the following holds.

$$p(\tilde{z}) \geq p(z) - \lambda^{*\top}(\tilde{z} - z)$$

To show that $-\lambda^*$ is a subgradient of $p$ at $z$, the value of $p$ is considered at another point $\tilde{z}$ [1]:

$$\begin{aligned} p(\tilde{z}) &= \sup_{\lambda \geq 0} \inf_{\theta \in \Theta} \left( f(\theta) + \lambda^\top (h(\theta) - \tilde{z}) \right) \\ &\geq \inf_{\theta \in \Theta} \left( f(\theta) + \lambda^{*\top} (h(\theta) - \tilde{z}) \right) \\ &= \inf_{\theta \in \Theta} \left( f(\theta) + \lambda^{*\top} (h(\theta) - z + z - \tilde{z}) \right) \\ &= \inf_{\theta \in \Theta} \left( f(\theta) + \lambda^{*\top} (h(\theta) - z) \right) + \lambda^{*\top}(z - \tilde{z}) \\ &= \phi(z) + (-\lambda^*)^\top (\tilde{z} - z) \end{aligned}$$

This holds for all $\tilde{z}$ in the domain of $p$. So we can conclude $-\lambda^*$ is a subgradient of $p$ at $z$.

# Problem 2

The combined consensus/projected incremental subgradient method for $N$ agents is considered:

$$\theta_{k+1}^{i} = \mathcal{P}_{\Theta}\left[\sum_{j=1}^{N}[W^{\varphi}]_{ij}(\theta_k^j - \alpha_k g^j(\theta_k^j))\right], \quad i = 1, ..., N \tag{1}$$

where $\theta_k^i$ is agents $i$ local variable at step $k$, $\varphi$ is the number of consensus steps that each agents runs with its neighbors, $W$ the communication graph, $\alpha_k$ the step size, $g^j(\theta_k^j)$ the subgradients of $f^j$ at $\theta_k^j$ ($f^j : \mathbb{R}^M \to \mathbb{R}$) and $\mathcal{P}$ denotes the projection on set $\Theta$. The communication graph is assumed to be strongly connected, balanced and doubly stochastic. To show that as $\varphi \to \infty$, the combined consensus/projected incremental subgradient method becomes a standard subgradient method, [2] was used. The standard subgradient method is given by:

$$\theta_{k+1} = \mathcal{P}_{\Theta}\left[\theta_k + \alpha_k \sum_{i=1}^{N} g^i(\theta_k)\right] \tag{2}$$

We assume that if in each subgradient iteration $\varphi \to \infty$, and that the consensus update converges to the average of the initial values. This holds since the communication graph is double stochastic (i.e. all rows and columns add up to 1) and $W = W^{\top}$. The for all $i = 1, ..., N$ and $\theta_0^i \in \mathbb{R}^M$:

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

Let us denote the initial state of the projected consensus variable with:

$$\breve{\theta}_0 = \mathcal{P}_{\Theta}\left[\frac{1}{N}\sum_{j=1}^{N}(\theta_0^j - \alpha_0 g^j(\theta_0^j)\right]$$

In the next iteration, each agent $i$ will possess the same value $\breve{\theta}_1$, thus the local subgradients $g^j$ will be evaluated at the same point. For $k \geq 1$, equation 1 is equivalent to:

$$\begin{aligned}\breve{\theta}_{k+1} &= \mathcal{P}_{\Theta}\left[\frac{1}{N}\sum_{j=1}^{N}(\breve{\theta}_k - \alpha_k g^j(\breve{\theta}_k)\right] \\ &= \mathcal{P}_{\Theta}\left[\breve{\theta}_k - \alpha_k \frac{1}{N}\sum_{j=1}^{N}(g^j(\breve{\theta}_k)\right]\end{aligned}$$

This is the same process as the standard subgradient method in equation 2.

# Problem 3

Each aircraft $i$ can be described by a linear time-invariant dynamic system:

$$x_i(t+1) = A_i x_i(t) + B_i u_i(t), \quad x_i(0) = x_{i,0}, \quad i = 1, ..., 4, \quad t = 0, ..., T_{final} - 1 \tag{3}$$

The cost function for each aircraft to be minimized is:

$$J_i(\mathbf{x_i}, \mathbf{u_i}) = \sum_t x_i(t)^\top x_i(t) + u_i(t)^\top u_i(t)$$

With equation 3 the cost function becomes:

$$\begin{aligned} J_i(x_{i,0}, \mathbf{u_i}) &= \left\| F_i x_{i,0} + \Phi_i \mathbf{u_i} \right\|^2 + \|\mathbf{u_i}\|^2 \\ &= \mathbf{u_i}^\top (\Phi_i^\top \Phi_i + \mathbf{I}_8)\mathbf{u_i} + 2x_{i,0}^\top F_i^\top \Phi_i \mathbf{u_i} + x_{i,0}^\top F_i^\top F_i x_{i,0} \end{aligned}$$

$$F_i = \begin{bmatrix} \mathbf{I}_4 \\ A_i \\ A_i^2 \\ A_i^3 \\ A_i^4 \\ A_i^5 \end{bmatrix} \quad \Phi_i = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ B_i & 0 & 0 & 0 & 0 \\ A_i B_i & B_i & 0 & 0 & 0 \\ A_i^2 B_i & A_i B_i & B_i & 0 & 0 \\ A_i^3 B_i & A_i^2 B_i & A_i B_i & B_i & 0 \\ A_i^4 B_i & A_i^3 B_i & A_i^2 B_i & A_i B_i & B_i \end{bmatrix} \quad \mathbf{u_i} = \begin{bmatrix} u_i(0) \\ u_i(1) \\ u_i(2) \\ u_i(3) \\ u_i(4) \end{bmatrix}$$

The minimization problem becomes:

$$\begin{aligned} & \min_{\mathbf{u_i}} \sum_i J_i(x_{i,0}, \mathbf{u_i}) \\ \text{subjected to} \quad & x_i(5) = x_f \\ & \mathbf{u_i}^\top \mathbf{u_i} \leq u_{max} \end{aligned} \tag{4}$$

with $x_f \in \mathbb{R}^4$ and $u_{max} = 100$. The first constraint can be rewritten as three linear constraints, which will function as the coupling constraints:

$$\begin{aligned} x_1(5) - x_2(5) &= 0 \\ x_2(5) - x_3(5) &= 0 \\ x_3(5) - x_4(5) &= 0 \end{aligned}$$

$$x_i(5) = A_i^5 x_{i,0} + A_i^4 B_i u_i(0) + A_i^3 B_i u_i(1) + A_i^2 B_i u_i(2) + A_i B_i u_i(3) + B_i u_i(4)$$

The three coupling constraints can be captured in three Lagrange multipliers $\lambda_i$. The partial Lagrangian for the dual problem is:

$$\begin{aligned} L(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}, \lambda) = {} & J_1(x_{1,0}, \mathbf{u_1}) + J_2(x_{2,0}, \mathbf{u_2}) + J_3(x_{3,0}, \mathbf{u_3}) + J_4(x_{4,0}, \mathbf{u_4}) \\ & + \lambda_1^\top (x_1(5) - x_2(5)) + \lambda_2^\top (x_2(5) - x_3(5)) + \lambda_3^\top (x_3(5) - x_4(5)) \end{aligned}$$

The dual problem can be solved in parallel, followed by an update of the Lagrange multipliers. The update rule for the Lagrange multipliers is:

$$\begin{aligned} \lambda_1(k+1) &= \lambda_1(k) + \alpha(x_1(5) - x_2(5)); \\ \lambda_2(k+1) &= \lambda_2(k) + \alpha(x_2(5) - x_3(5)); \\ \lambda_3(k+1) &= \lambda_3(k) + \alpha(x_3(5) - x_4(5)); \end{aligned}$$

where $\alpha$ is the step size of the update rule. If $\alpha$ is chosen too small, the solution might converge very slow and if $\alpha$ is chosen too big the solution might not converge at all. First, $\alpha$ was chosen as 0.001, but then it would take the problem more than 30000 iterations to converge. So $\alpha$ was increased to 0.01 such that the solution would converge faster. The number of iterations needed with $\alpha = 0.01$ is 11484. The algorithm is stopped when the error between the four states is smaller than 0.001. The final state is $\mathbf{x_i}(5) = \begin{bmatrix} -1.95 & -1.62 & -0.82 & -0.97 \end{bmatrix}^\top$ and the convergence is show in Figure 1.

The quadratic constraint on the input in Equation 4 is rewritten as a set of linear constraints:

$$\begin{aligned} \begin{bmatrix} 1 & 1 \end{bmatrix} \otimes \mathbf{I_5} &\leq 20 \\ \begin{bmatrix} -1 & -1 \end{bmatrix} \otimes \mathbf{I_5} &\leq 20 \\ \begin{bmatrix} 1 & -1 \end{bmatrix} \otimes \mathbf{I_5} &\leq 20 \\ \begin{bmatrix} -1 & 1 \end{bmatrix} \otimes \mathbf{I_5} &\leq 20 \end{aligned}$$

where $\otimes$ is the Kronecker product.

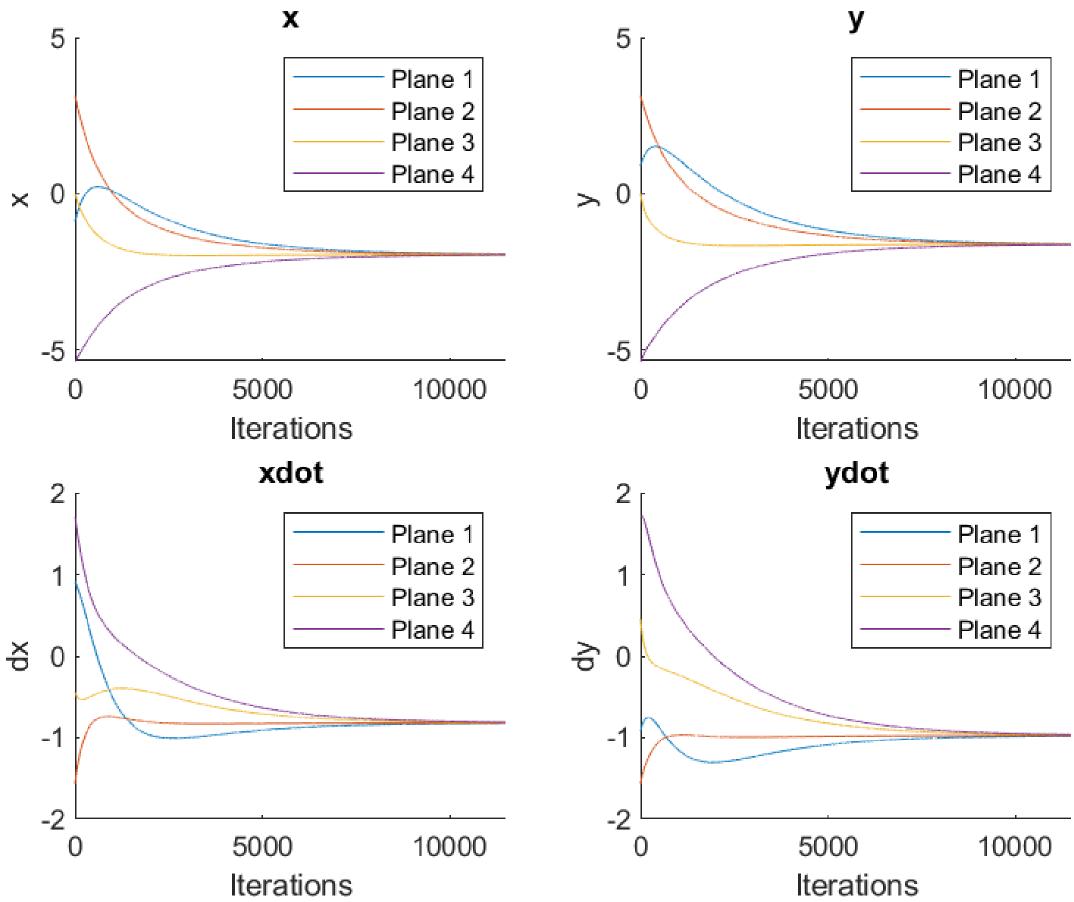


Figure 1: Convergence of the dual problem.

# Exercise 4

a) The derivative of the cost function $V(u_1, u_2)$ with respect to one of the variables $u_i$ is taken while keeping the other fixed, to show that the next iteration of the optimization procedure is

$$u_1^{p+1} = -H_{11}^{-1}\left(H_{12}u_2^p + c_1\right)$$
$$u_2^{p+1} = -H_{22}^{-1}\left(H_{21}u_1^p + c_2\right)$$

The cost function, written out in terms of $u_1$ and $u_2$, is given by

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

Taking the partial derivative with respect to $u_1$ and to $u_2$ and using the fact that $H_{11}^\top = H_{11}$, $H_{22}^\top = H_{22}$ and $H_{12} = H_{21}^\top$:

$$\frac{\partial V(u_1, u_2)}{\partial u_1^\top} = \frac{1}{2}H_{11}u_1 + \frac{1}{2}H_{11}^\top u_1 + \frac{1}{2}H_{21}^\top u_2 + \frac{1}{2}H_{21}u_2 + c_1 = H_{11}u_1 + H_{12}u_2 + c_1$$

$$\frac{\partial V(u_1, u_2)}{\partial u_2^\top} = \frac{1}{2}H_{21}u_1 + \frac{1}{2}H_{12}^\top u_1 + \frac{1}{2}H_{22}u_2 + \frac{1}{2}H_{22}^\top u_2 + c_2 = H_{22}u_2 + H_{21}u_1 + C_2$$

Setting these partial derivatives equal to zero and rewriting, this yields the next iteration for $u_1$ and $u_2$:

$$u_1^{p+1} = -H_{11}^{-1}\left(H_{21}u_2^p + c_1\right)$$

$$u_2^{p_2} = -H_{22}^{-1}\left(H_{21}u_1^p + c_2\right)$$

b) To show that the iteration matrix is stable, $|\text{eig}(A)| < 1$, it needs to be shown that the eigenvalues of $A$ are within the interval (-1,1).

When $\text{eig}(A + I) > 0$, and $\text{eig}(A - I) < 0$, it follows that $\text{eig}(A) \in (-1, 1)$ since the following property holds for matrices

$$\text{eig}(A + I) = \text{eig}(A) + \text{eig}(I) = \text{eig}(A) + I$$
$$\text{eig}(A - I) = \text{eig}(A) - \text{eig}(I) = \text{eig}(A) - I$$

Here, the LU decomposition of $A + I$ is

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

So, when $\text{eig}(I - H_{11}^{-1}H_{12}H_{22}^{-1}H_{21}) > 0 \implies \text{eig}(A + I) > 0$.

It is known that $H > 0$ and $H_{22}$ too:

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} > 0$$

To prove that $\text{eig}(I - H_{11}^{-1}H_{12}H_{22}^{-1}H_{21}) > 0$, the Schur complement of the $H_{22}$ block in the $H$ matrix is introduced, where $H/H_{22} > 0$ since both $H$ and $H_{22}$ are positive definite:

$$H/H_{22} = H_{11} - H_{12}H_{22}^{-1}H_{21} > 0$$
$$H_{11}^{-1}H_{11} - H_{11}^{-1}H_{12}H_{22}^{-1}H_{21} > 0$$
$$I - H_{11}^{-1}H_{12}H_{22}^{-1}H_{21} > 0$$

Combining the above, it is shown that $\text{eig}(A) > -1$:

$$
\begin{aligned}
I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{21} &> 0 \\
\mathrm{eig}(I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{21}) &> 0 \\
\mathrm{eig}(A + I) &> 0 \\
\mathrm{eig}(A) + I &> 0 \\
\mathrm{eig}(A) &> -1
\end{aligned}
$$

The same procedure can be applied to show that $\mathrm{eig}(A)$ do not exceed the upper limit of the interval $(-1, 1)$. $|\mathrm{eig}(A)| < 1$.

c) Because $|\mathrm{eig}(A)| < 1$ as shown in the previous question, powers of A converge to 0 in the limit. The limit of $u^p$ is given by the following expression, where the last term is rewritten using the geometric series for matrices.

$$
\lim_{p\to\infty} u^{p+1} = A^{\infty} u^p + \sum_{p=0}^{\infty} (A^p)\, b = \sum_{p=0}^{\infty} (A^p)\, b = (I - A)^{-1} b \tag{5}
$$

By rewriting the expression of $(I - A)^{-1} b$ as follows, it is shown that the iteration produces the same solution as $u^* = -H^{-1} c$:

$$
\begin{aligned}
u^* = (I - A)^{-1} b = &\begin{bmatrix} I & H_{11}^{-1} H_{12} \\ H_{22}^{-1} H_{12}^\top & I \end{bmatrix} b \\
= &\begin{bmatrix} -(I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} & (I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{11}^{-1} H_{12} & -H_{11}^{-1} & 0 \\ H_{22}^{-1} H_{12}^\top (I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} & -H_{22}^{-1} - H_{22}^{-1} H_{12}^\top (I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{11}^{-1} H_{12} & 0 & -H_{22}^{-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
= &\begin{bmatrix} -(I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{11}^{-1} & (I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{11}^{-1} H_{12} H_{22}^{-1} \\ H_{22}^{-1} H_{12}^\top (I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{11}^{-1} & -H_{22}^{-1} - H_{22}^{-1} H_{12}^\top (I - H_{11}^{-1} H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{11}^{-1} H_{12} H_{22}^{-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
= -&\begin{bmatrix} (H_{11} - H_{12} H_{22}^{-1} H_{12}^\top)^{-1} & -(H_{11} - H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{12} H_{22}^{-1} \\ H_{22}^{-1} H_{12}^\top (H_{11} - H_{12} H_{22}^{-1} H_{12}^\top)^{-1} & -H_{22}^{-1} + H_{22}^{-1} H_{12}^\top (H_{11} - H_{12} H_{22}^{-1} H_{12}^\top)^{-1} H_{12} H_{22}^{-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\
= -&\begin{bmatrix} H_{11} & H_{12} \\ H_{12}^\top & H_{22} \end{bmatrix}^{-1} c \\
= -&H^{-1} c = u^*
\end{aligned}
$$

# Exercise 5

To prove that the cost function $V$ is monotonically decreasing, it needs to be proven that:

$$V(u^{p+1}) < V(u^p) \quad \forall u^p \neq -H^{-1}c \tag{6}$$

Before this is proven, it will first be shown that:

$$V(u^{p+1}) - V(u^p) = -\frac{1}{2}(u^p - u^*)^T P (u^p - u^*)$$

To do so, the coordinate system is moved to the optimal value $u^*$. By making use of $u^* = Au^* + b$ since the iteration converges, this results in the coordinates:

$$\begin{aligned}
\tilde{u}^p &= u^p - u^* \\
\tilde{u}^{p+1} &= u^{p+1} - u^* \\
&= Au^p + b - u^* \\
&= A\tilde{u}^p + Au^* + b - u^* \\
&= A\tilde{u}^p
\end{aligned}$$

Equation 6 becomes:

$$\begin{aligned}
\frac{1}{2}\tilde{u}^{p\top} A^\top HA\tilde{u}^p + c^\top A\tilde{u}^p - \frac{1}{2}\tilde{u}^{p\top} H\tilde{u}^p + c^\top \tilde{u}^p \leq 0 \\
\frac{1}{2}\tilde{u}^{p\top}(A^\top HA - H)\tilde{u}^p + c^\top A\tilde{u}^p - c^\top \tilde{u}^p \leq 0
\end{aligned} \tag{7}$$

Since in the changed coordinate system, $u^*$ is at the origin, $u^* = 0 = -H^{-1}c \quad \longrightarrow \quad c = 0$. Using this, Equation 7 becomes

$$\frac{1}{2}\tilde{u}^{p\top}(A^\top HA - H)\tilde{u}^p \leq 0$$

To prove that $A^\top HA - H = -P$, the following equalities are used:

$$\begin{aligned}
H &= D + N \\
A &= -D^{-1}N
\end{aligned}$$

Then, P is rewritten to:

$$\begin{aligned}
P &= HD^{-1}\tilde{H}D^{-1}H = (D+N)D^{-1}(D-N)D^{-1}(D+N) \\
P &= (I + ND^{-1})(I - ND^{-1})(D+N) \\
P &= (I - ND^{-1} + ND^{-1} - ND^{-1}ND^{-1})(D+N) \\
P &= D + N - ND^{-1}N - ND^{-1}ND^{-1}N
\end{aligned}$$

Then, $A^\top HA - H$ is rewritten to P as follows:

$$\begin{aligned}
A^\top HA - H &= ND^{-1}(D+N)D^{-1}N - (D+N) \\
&= -D - N + ND^{-1}N + ND^{-1}ND^{-1}N \\
&= -P
\end{aligned}$$

To show that Equation 6 is monotonically decreasing, it needs to be proven that P is positive definite, since then

$$V(u^{p+1}) - V(u^p) = -\frac{1}{2}(u^p - u^*)^T P (u^p - u^*) \leq 0$$

Since $P = HD^{-1}\tilde{H}D^{-1}H$, $D^{-1}H$ nonsingular and

$$\tilde{H} = D - N = \begin{bmatrix} H_{11} & 0 \\ 0 & H_{22} \end{bmatrix} - \begin{bmatrix} 0 & H_{12} \\ H_{21} & 0 \end{bmatrix} = \begin{bmatrix} H_{11} & -H_{12} \\ -H_{21} & H_{22} \end{bmatrix} > 0$$

Then, P is in the form of $R^\top QR$ with $Q = \tilde{H}$ and $R = D^{-1}H$ nonsingular, thus $R^\top QR > 0$, so $P > 0$. The cost function is monotonically decreasing.

# References

[1] S. Boyd, L. Xiao, A. Mutapic, and J. Mattingley. Notes on decomposition methods for ee364b. *Stanford University*, Winter 2006-07.

[2] B. Johansson, T. Keviczky, M. Johansson, and K. H. Johansson. Subgradient methods and consensus algorithms for solving convex optimization problems. *2008 47th IEEE Conference on Decision and Control*, pages 4185–4190, Dec 2008.