Assignment 3

Networked and Distributed Control - SC42101

J.C.S. Amato (4688368)
TU Delft
Delft Center for Systems and Control

June 11, 2024

1 Primal and Dual Decomposition

1.a Primal Decomposition

We consider the following convex optimization problem with complicating constraint:

$$\begin{aligned} \min_{x_1,x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{subject to} \quad & x_1 \in \mathcal{X}_1, \quad x_2 \in \mathcal{X}_2 \\ & h_1(x_1) + h_2(x_2) \geq 0 \end{aligned} \tag{1.1}$$

where $\mathcal{X}_1, \mathcal{X}_2$ are convex sets and all functions are convex. To apply primal decomposition, we first rewrite the problem by introducing a coupling variable r:

$$\begin{aligned} \min_{x_1,x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{subject to} \quad & x_1 \in \mathcal{X}_1, \quad x_2 \in \mathcal{X}_2 \\ & h_1(x_1) \leq r, \quad h_2(x_2) \leq -r \end{aligned} \tag{1.2}$$

Such that:

$$h_1(x_1) + h_2(x_2) \leq r - r = 0$$

Hence, we can solve two subproblems at the lower level, given a value for r. **Subproblem 1** becomes:

$$\begin{aligned} \min_{x_1} \quad & f_1(x_1) \\ \text{subject to} \quad & x_1 \in \mathcal{X}_1 \\ & h_1(x_1) \leq r \end{aligned} \tag{1.3}$$

Similarly, we have for **Subproblem 2**:

$$\begin{aligned} \min_{x_2} \quad & f_2(x_2) \\ \text{subject to} \quad & x_2 \in \mathcal{X}_2 \\ & h_2(x_2) \leq -r \end{aligned} \tag{1.4}$$

Subsequently, we eliminate the primal variables. Let $\phi_1(r)$ and $\phi_2(r)$ denote the optimal value of the subproblems 1 and 2 respectively, given the constraints. That is:

$$\begin{aligned} \phi_1(r) &= \inf_{x_1}\{f_1(x_1) \mid h_1(x_1) \leq r,\ x_1 \in \mathcal{X}_1\} \\ \phi_2(r) &= \inf_{x_2}\{f_2(x_2) \mid h_2(x_2) \leq -r,\ x_2 \in \mathcal{X}_2\} \end{aligned} \tag{1.5}$$

Such that we can solve the **Master problem** at the higher level:

$$\begin{aligned} &\min_{r} \phi_1(r) + \phi_2(r) \\ &\text{subject to} \quad r \in \mathbb{R}\backslash\{-\infty, +\infty\} \end{aligned} \tag{1.6}$$

Therefore, the role of the master problem is to update the coupling variable r in the respective subproblems by solving a parametric optimization problem.

1.b Dual Decomposition

Let us first rewrite the optimization problem in the standard form:

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & x \in \mathcal{X} \\ & h(x) - z \leq 0 \end{aligned} \tag{1.7}$$

such that the Lagrangian becomes:

$$L(x, \lambda, z) = f(x) + \lambda^T (h(x) - z) \tag{1.8}$$

Note that the Lagrangian function explicitly depends on z. To find the optimal dual function, we maximize the Lagrangian over the dual variable λ and minimize it over the primal variable x. This process yields the optimal dual function $d^*(z)$, defined as:

$$d^*(z) = \sup_{\lambda \geq 0} \inf_x L(x, \lambda, z) \tag{1.9}$$

For the subsequent analysis, it is important to state the definition of a subgradient.

Definition 1 *g is a subgradient for f at x_0 if $g \in \partial f(x_0)$, where $\partial f(x_0) := \{g \in \mathbb{R}^M | f(x) \geq f(x_0) + g^T(x - x_0)$.*

Furthermore, let us assume that $\exists\ x \in relint(\mathcal{X})$ such that $h(x) - z < 0$. Then, since the problem is (strongly) convex, this implies that Slaters condition holds. Slaters condition is a sufficient condition for strong duality. Let $\lambda^*(z)$ be the optimal dual variable corresponding to the inequality constraint in the primal problem $h(x) \leq z$, then:

$$\begin{aligned} p(z) = d^*(z) &= \sup_{\lambda \geq 0} \inf_x L(x, \lambda, z) \\ &= \inf_x [f(x) + \lambda^*(z)^T(h(x) - z)] \\ &\leq f(x) + \lambda^*(z)^T(h(x) - z) \end{aligned} \tag{1.10}$$

Where the latter inequality holds for any $x \in \mathcal{X}$, by definition of the infimum. If we consider a different constraint $h(x) - z \leq u$, then we become:

$$p(z) \leq f(x) + \lambda^*(z)^T(h(x) - z) \leq f(x) + \lambda^*(z)^T u \tag{1.11}$$

Let us redefine the constraint in standard form:

$$h(x) - z - u = h(x) - \tilde{z} \leq 0$$

Then observe that $p(\tilde{z})$ is nothing else than $f(x^*, \tilde{z})$, i.e. the objective function evaluated in the optimizer $x^* \in \mathcal{X}$ given the constraint $h(x) - \tilde{z} \leq 0$. Hence, we can rewrite the latter inequality in equation 1.11 as:

$$\begin{aligned} p(z) &\leq p(\tilde{z}) + \lambda^*(z)^T u \\ p(\tilde{z}) &\geq p(z) - \lambda^*(z)^T u \end{aligned} \tag{1.12}$$

Then, since $\tilde{z} = z + u \implies \tilde{z} - z = u$ such that $h(x) - \tilde{z} = h(x) - z - u$:

$$p(\tilde{z}) \geq p(z) - \lambda^*(z)^T(\tilde{z} - z) \tag{1.13}$$

In conclusion, by definition 1 it follows that $-\lambda(z)^*$ is a subgradient of p at z.

2 Consensus Iterations

At first, some useful facts are listed that follow from the assumptions on the matrix W. To find the limit of W, we use Lemma 1 (see Lemma 2.3.3 in Johannson [2]):

Lemma 1 *If and only if Assumption 1 b)-d) is fulfilled, then:*

$$\lim_{k \to \infty} W^k = \frac{\mathbf{1}_n \mathbf{1}_n^T}{n} \quad \textit{and} \quad \lim_{k \to \infty} y^{(k)} = \bar{y}^{(0)}.$$

The assumptions referenced in the Lemma are given below (see Assumption 2.3.2 in Johannson [2]).

Assumption 1 *The matrix $W \in \mathbb{R}^{n \times n}$, with $n = |\mathcal{V}|$, fulfills*

(a) $[W]_{ij} = 0$, *if* $(i, j) \notin \mathcal{E}$ *and* $i \neq j$,

(b) $\mathbf{1}_n^T W = \mathbf{1}_n^T$,

(c) $W \mathbf{1}_n = \mathbf{1}_n$,

(d) $\rho(W - \mathbf{1}_n \mathbf{1}_n^T / n) \leq \gamma < 1$,

where $\mathbf{1}_n \in \mathbb{R}^n$ is the column vector with all elements equal to one and $\rho(\cdot)$ is the spectral radius.

Indeed, since $W \mathbf{1}_N = \mathbf{1}_N$ we have that by symmetry of W, $\mathbf{1}_n^T W = \mathbf{1}_n^T$ hence assumption 1 (b) holds. In addition, assumption 1 (d) holds as given in the problem statement.

Proof of Lemma 1. Given that $\rho(W - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N}) < 1$, we aim to show that $\lim_{t \to \infty} (W - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N})^t = 0$. First, we diagonalize the matrix power $(W - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N})^t$. Let Q be the matrix of eigenvectors and Λ be the diagonal matrix of eigenvalues, such that:

$$(W - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N})^t = (Q^{-1} \Lambda Q)^t = Q^{-1} \Lambda^t Q$$

Since the eigenvalues $\lambda \in \Lambda$ satisfy $|\lambda| < 1$, it follows that $\lim_{t \to \infty} \lambda^t = 0$. Therefore, Λ^t becomes the zero matrix in the limit. Thus, we can rewrite:

$$\begin{aligned} (W - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N})^t &= (W(I - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N}))^t \\ &= W^t (I - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N})^t \\ &= W^t (I - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N}) \\ &= W^t - \frac{\mathbf{1}_N \mathbf{1}_N^T}{N} \end{aligned}$$

The first equality follows from assumption 1 (c), that is $W\mathbf{1}_N = \mathbf{1}_N$ implies $W\mathbf{1}_N\mathbf{1}_N^T = \mathbf{1}_N\mathbf{1}_N^T$. The former equality follows from the fact that $I - \frac{\mathbf{1}_N\mathbf{1}_N^T}{N}$ is a projection matrix. Hence,by substitution:

$$\lim_{t\to\infty}(W - \frac{\mathbf{1}_N\mathbf{1}_N^T}{N})^t = \lim_{t\to\infty} W^t - \frac{\mathbf{1}_N\mathbf{1}_N^T}{N} = 0$$

$$\lim_{t\to\infty} W^t = \frac{\mathbf{1}_N\mathbf{1}_N^T}{N}$$

Which ends the proof of Lemma 1 ■

For further analysis we firstly introduce for a more compact notation:

$$z_k^i = x_k^i - \alpha_k g^i(x_k^i) \tag{2.1}$$

Then we make use of lemma 1 to observe that for all $i = 1, \dots, N$ and $x_0^i \in \mathbb{R}^N$:

$$\begin{aligned}\lim_{\phi\to\infty}\sum_{j=1}^{N} W_{ij}^{\phi} z_0^j &= \frac{1}{N}(z_0^j + z_0^{j+1} + \ldots + z_0^N) \\ &= \frac{1}{N}\sum_{j=1}^{N}(x_0^j - \alpha_0 g^j(x_0^j))\end{aligned} \tag{2.2}$$

Consequently, let us then denote the first iterate of the projected consensus variable by:

$$\bar{x}_1^i = \mathcal{P}_{\mathcal{X}}\left[\frac{1}{N}\sum_{j=1}^{N} x_0^j - \alpha_0 g^j(x_0^j)\right],\ i = 1, \dots, N \tag{2.3}$$

Hence, in the next iteration each agent will have the same iterate $\bar{x}_1$. In addition, all local subgradients g_j are evaluated at the same point leading to the update equation for $k \geq 1$:

$$\bar{x}_{k+1} = \mathcal{P}_{\mathcal{X}}\left[\bar{x}_k - \alpha_k \frac{1}{N}\sum_{j=1}^{N} g^j(\bar{x}_k^j)\right],\ i = 1, \dots, N \tag{2.4}$$

Which is the procedure for the standard subgradient method with a step size of α_k/N. Therefore, we can conclude that when $\phi \to \infty$ the combined consensus/projected incremental subgradient method becomes a standard subgradient method.

3 Componentwise optimisation

3.a Preliminaries

In this subsection, we first derive the general analytic solution to the optimization problem in u. We then present several useful facts that will be essential for the derivation in Section 3.a.1.

Consider the objective function $V(u) = \frac{1}{2}u^T Hu + c^T u + d$ with $H \succ 0$. To find the minimizer, we compute the gradient of $V(u)$:

$$\nabla V(u) = u^T H + c^T \tag{3.1}$$

Setting the gradient to zero gives us the first-order optimality condition. Solving for u, we get the unique analytic solution:

$$u^* = -H^{-1}c \tag{3.2}$$

The solution u^* is unique because $H \succ 0$ ensures that H is invertible and the problem is strictly convex.

$H \succ 0$ implies $H = H^T$, hence we note that $H_{12} = H_{21}^T$. However, $H \succ 0$ does not directly imply that H_{11} and H_{22} are positive definite or invertible. In fact, by the Schur complement (see proposition 1), we can see that if either one of the diagonal blocks is invertible then $H \succ 0$ if and only if the block itself and the Schur complement with respect to that block are positive definite. Therefore, for this procedure, we must assume that H_{11} and H_{22} are non-singular which by proposition 1 implies that $H_{11} \succ 0$ and $H_{22} \succ 0$.

Proposition 1 *For any symmetric matrix, M, of the form*

$$M = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix},$$

if A is invertible then the following properties hold:

1. *$M \succ 0$ if and only if $A \succ 0$ and $C - B^\top A^{-1}B \succ 0$.*
2. *If $A \succ 0$, then $M \succeq 0$ if and only if $C - B^\top A^{-1}B \succeq 0$.*

if C is invertible then the following properties hold:

1. *$M \succ 0$ if and only if $C \succ 0$ and $A - B^\top C^{-1}B \succ 0$.*
2. *If $A \succ 0$, then $M \succeq 0$ if and only if $A - B^\top C^{-1}B \succeq 0$.*

3.A.1 Optimization Procedure as a Dynamical System

The objective function may be written out more explicitly as:

$$\begin{aligned} V(u) &= \frac{1}{2}u^T H u + c^T u + d \\ &= \frac{1}{2}(u_1^T H_{11} u_1 + u_2^T H_{22} u_2 + u_1^T H_{12} u_2 \\ &\quad + u_2^T H_{21} u_1) + c_1^T u_1 + c_2^T u_2 + d \end{aligned} \tag{3.3}$$

From which we obtain the partial derivatives in u_1 and u_2:

$$\begin{aligned} \frac{\partial V}{\partial u_1} &= u_1^T H_{11} + \frac{1}{2} u_2^T H_{12}^T + \frac{1}{2} u_2^T H_{21} + c_1^T \\ \frac{\partial V}{\partial u_2} &= u_2^T H_{22} + \frac{1}{2} u_1^T H_{12} + \frac{1}{2} u_1^T H_{21}^T + c_2^T \end{aligned} \tag{3.4}$$

Setting $\frac{\partial V}{\partial u_1} = 0$ we obtain:

$$\begin{aligned} u_1^T H_{11} + u_2^T H_{12}^T + c_1^T &= 0 \\ u_1^T H_{11} &= -u_2^T H_{12}^T - c_1^T \\ u_1^T &= -(u_2^T H_{12}^T + c_1^T) H_{11}^{-1} \\ u_1 &= -H_{11}^{-T}(H_{12} u_2 + c_1) \\ u_1 &= -H_{11}^{-1}(H_{12} u_2 + c_1) \end{aligned} \tag{3.5}$$

The third equality comes by the non-singularity assumption on both H_{11} and H_{22} and the last equality comes from $H_{11} \succ 0 \implies H_{11}^T = H_{11}$. Similarly, we have for $\frac{\partial V}{\partial u_2} = 0$:

$$\begin{aligned} u_2^T H_{22} + u_1^T H_{12} + c_2^T &= 0 \\ u_2^T &= -(u_1^T H_{12} + c_2^T) H_{22}^{-1} \\ u_2 &= -H_{22}^{-1}(H_{12} u_1 + c_2) \end{aligned} \tag{3.6}$$

Summarizing, if we take an initial point (u_1^p, u_2^p), then:

$$\begin{aligned} u_1^{p+1} &= -H_{11}^{-1}(H_{12} u_2^p + c_1) \\ &= -H_{11}^{-1} H_{12} u_2^p - H_{11}^{-1} c_1 \\ u_2^{p+1} &= -H_{22}^{-1}(H_{12} u_1^p + c_2) \\ &= -H_{22}^{-1} H_{12} u_1^p - H_{22}^{-1} c_2 \end{aligned} \tag{3.7}$$

Which is in fact a discrete dynamical system:

$$u^{p+1} = A u^p + b \tag{3.8}$$

where:

$$A = \begin{pmatrix} 0 & -H_{11}^{-1} H_{12} \\ -H_{22}^{-1} H_{12} & 0 \end{pmatrix}, b = \begin{pmatrix} -H_{11}^{-1} c_1 \\ -H_{22}^{-1} c_2 \end{pmatrix} \tag{3.9}$$

3.B Convergence Analysis

To prove convergence of the algorithm we make use of the discrete dynamical system representation of the iteration procedure. That is, the iteration procedure is asymptotically 'stable' if and only if:

$$|eig(A)| < 1 \tag{3.10}$$

To prove this property, we use the *Householder-John* Theorem, stated in theorem 1 [1].

Theorem 1 *If M and L are real matrices such that both (a) M and (b) $M - L - L^T$ are symmetric positive definite, then the spectral radius of $-(M - L)^{-1}L$ is strictly less than one.*

For this particular problem, we can define the following matrices:

$$M = H \succ 0, \quad L = \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix} \tag{3.11}$$

By symmetry of the antidiagonal blocks $L = L^T$ hence:

$$M - L - L^T = M = \begin{pmatrix} H_{11} & -H_{12} \\ -H_{21} & H_{22} \end{pmatrix} \succ 0 \tag{3.12}$$

which verifies that assumption (a) and (b) of theorem 1 are satisfied. Note that the positive definiteness of the matrix follows by the useful facts stated in the assignment. Then, observe that:

$$\begin{aligned} M - L &= \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix} \\ &\Leftrightarrow \\ (M - L)^{-1} &= \begin{pmatrix} H_{11}^{-1} & 0 \\ 0 & H_{22}^{-1} \end{pmatrix} \\ &\Leftrightarrow \\ -(M - L)^{-1} L &= A \end{aligned} \tag{3.13}$$

Therefore by theorem 1:

$$eig(|-(M - L)^{-1} L|) = eig(|A|) < 1 \tag{3.14}$$

3.C Comparison to analytic solution

The analytic solution to the minimization problem:

$$u^* = -H^{-1} c = \begin{pmatrix} u_1^* \\ u_2^* \end{pmatrix} = -\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}^{-1} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{3.15}$$

can be rewritten as:

$$
\begin{aligned}
Hu^* &= -c \\
Hu^* &= \begin{pmatrix} H_{11}u_1^* + H_{12}u_2^* \\ H_{21}u_1^* + H_{22}u_2^* \end{pmatrix} = -\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}
\end{aligned} \tag{3.16}
$$

We proceed to separate the equations, leading to:

$$
\begin{aligned}
H_{11}u_1^* + H_{12}u_2^* &= -c_1 \\
H_{21}u_1^* + H_{22}u_2^* &= -c_2
\end{aligned} \tag{3.17}
$$

which leads to the final form:

$$
\begin{aligned}
u_1^* &= -H_{11}^{-1}(H_{12}u_2^* + c_1) \\
u_2^* &= -H_{22}^{-1}(H_{21}u_1^* + c_2)
\end{aligned} \tag{3.18}
$$

Subsequently, we must show that the procedure we have previously analyzed produces the same solution. Observe that the value of the dynamical system as a function of p can be determined using (I use $u(p)$ instead of u^p to avoid confusion with powers):

$$
u(p) = A^p u(0) + \sum_{j=1}^{p-1} A^{p-1-j} b \tag{3.19}
$$

Note that since $|eig(A)| < 1$ then $\lim_{p\to\infty} A^p = 0$. Therefore, to find the convergence point we can simply observe that the second term is a geometric series with a known solution:

$$
\lim_{p\to\infty} u(p) = 0 + (I - A)^{-1} b = \overline{u} \tag{3.20}
$$

where we call $\overline{u}$ the unique solution to the componentwise procedure. Subsequently, we can proceed to work out the equations:

$$
\begin{aligned}
\overline{u} &= (I - A)^{-1} b \\
(I - A)\overline{u} &= b \\
\begin{pmatrix} I & H_{11}^{-1}H_{12} \\ H_{22}^{-1}H_{21} & I \end{pmatrix} \begin{pmatrix} \overline{u}_1 \\ \overline{u}_2 \end{pmatrix} &= \begin{pmatrix} -H_{11}^{-1}c_1 \\ -H_{22}^{-1}c_2 \end{pmatrix}
\end{aligned} \tag{3.21}
$$

Again, by separating the equations we obtain:

$$
\begin{aligned}
\overline{u}_1 &= -H_{11}^{-1}(H_{12}\overline{u}_2 + c_1) \\
\overline{u}_2 &= -H_{22}^{-1}(H_{21}\overline{u}_1 + c_2)
\end{aligned} \tag{3.22}
$$

Which is indeed the same solution as the analytic form, recalling that by strict convexity the objective function has a unique optimizer.

4 Decreasing Cost Function

In subsection 4.a it is firstly proven that the quadratic form with matrix P is monotonically decreasing. Subsequently, in section 4.b the closed form of the descent equation is derived as a function of H and A. Lastly, equivalence between the quadratic form with matrix P and the closed form in section 4.b is shown, by which we can conclude that the descent equation is monotonically decreasing.

4.a Positive definiteness of P

In this section we aim to prove that $P \succ 0$ such that the quadratic form $-\frac{1}{2}(u_p - u^*)^T P (u_p - u^*) \prec 0$. We first analyze the components of P:

$$
P = HD^{-1}\tilde{H}D^{-1}H \tag{4.1}
$$

The matrix is quadratic in HD^{-1}, which is equivalent to:

$$
\begin{aligned}
HD^{-1} &= \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} H_{11}^{-1} & 0 \\ 0 & H_{22}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} I & H_{11}^{-1}H_{12} \\ H_{22}^{-1}H_{21} & I \end{pmatrix} = (I - A)
\end{aligned} \tag{4.2}
$$

As shown in problem 3.c the unique optimizer of the componentwise iteration procedure exists and is given by $\bar{u} = (I - A)^{-1} b$. Therefore $(I - A)$ is non-singular. Furthermore:

$$
\tilde{H} = D - N = \begin{pmatrix} H_{11} & -H_{12} \\ -H_{21} & H_{22} \end{pmatrix} \succ 0 \tag{4.3}
$$

by the hint given in the assignment. We can then rewrite:

$$
P = (I - A)^T \tilde{H} (I - A) \tag{4.4}
$$

In conclusion, by the second hint in the assignment, since $\tilde{H}$ real, symmetric and positive definite, and $(I - A)$ real and nonsingular it follows that:

$$
P = (I - A)^T \tilde{H} (I - A) \succ 0 \tag{4.5}
$$

Hence the quadratic form in P:

$$
-\frac{1}{2}(u_p - u^*)^T P (u_p - u^*) \prec 0 \tag{4.6}
$$

which ends the proof.

4.b Descent Equation

The objective value function at iterate u^{p+1} must be firstly written out as a function of u_p, such that:

$$\begin{aligned} V(u^{p+1}) &= \frac{1}{2}(Au_p+b)^T H(Au_p+b) + c^T(Au_p+b) + d \\ &= \frac{1}{2}u_p^T A^T H A u_p u_p + \frac{1}{2}A^T u_p^T Hb + \frac{1}{2}b^T HAu_p \\ &\quad + \frac{1}{2}b^T Hb + c^T(Au_p+b) + d \end{aligned} \tag{4.7}$$

We can use this expression in the descent equation. The bias term d disappears and we can group some terms such that:

$$\begin{aligned} &V(u^{p+1}) - V(u_p) = \\ &\quad = \frac{1}{2}u_p^T(A^T HA - H)u_p + \left[\frac{1}{2}b^T HA + c^T(A-I)\right]u_p \\ &\quad + \frac{1}{2}b^T Hb + c^T b + \frac{1}{2}u_p^T A^T Hb \end{aligned} \tag{4.8}$$

Now notice that:

$$c^T(A-I)u_p = \frac{1}{2}c^T(A-I)u_p + \frac{1}{2}u_p^T(A^T-I)c \tag{4.9}$$

Whence we arrive at the final form where we define quadratic, linear and constant terms:

$$\begin{aligned} &V(u^{p+1}) - V(u_p) = \\ &\quad = \underbrace{\frac{1}{2}u_p^T(A^T HA - H)u_p}_{V_1} \\ &\quad + \frac{1}{2}\underbrace{\left[b^T HA + c^T(A-I)\right]}_{V_2}u_p \\ &\quad + \frac{1}{2}u_p^T\underbrace{\left[A^T Hb + (A^T-I)c\right]}_{V_3} \\ &\quad + \frac{1}{2}\underbrace{b^T Hb + c^T b}_{V_4} \end{aligned} \tag{4.10}$$

4.c Monotonically Decreasing Cost Function

The aim of this subsection is to prove that $V(u^{p+1}) - V(u_p) \prec 0$ by showing that $V(u^{p+1}) - V(u_p) = -\frac{1}{2}(u_p - u^*)^T P(u_p - u^*)$, such that the cost function is monotonically decreasing.

It is important to show some useful facts which will be used in the proof. Observe that, with D and N as defined in the assignment, by simple algebraic computation we have that:

$$\begin{aligned} D^{-1}H &= I + D^{-1}N \\ ND^{-1}H &= N(I + D^{-1}N) = (I + ND^{-1})N = HD^{-1}N \end{aligned} \tag{4.11}$$

Notice in the last equality that $D^{-1}H$ is symmetric since H is symmetric and D is symmetric by construction as it is a block diagonal matrix with symmetric elements.

In that context, we proceed to define the terms of the quadratic form in P:

$$\begin{aligned} &-\frac{1}{2}(u_p - u^*)^T P(u_p - u^*) \\ &= \underbrace{-\frac{1}{2}u_p^T P u_p}_{P_1} + \underbrace{\frac{1}{2}u_p^T P u^*}_{P_2} + \underbrace{\frac{1}{2}(u^*)^T P u_p}_{P_3} - \underbrace{\frac{1}{2}(u^*)^T P u^*}_{P_4} \end{aligned} \tag{4.12}$$

Let us decompose the term P:

$$\begin{aligned} &HD^{-1}\tilde{H}D^{-1}H = \\ &= HD^{-1}(D-N)D^{-1}H \\ &= HD^{-1}H - HD^{-1}ND^{-1}H \\ &= H(I + D^{-1}N) - (I + ND^{-1})HD^{-1}N \\ &= H + \underbrace{HD^{-1}N - ND^{-1}H}_{=0} + ND^{-1}HND^{-1} \end{aligned} \tag{4.13}$$

We use the fact that $HD^{-1}N = ND^{-1}H$. Finally, it can be seen that:

$$P = H + \underbrace{ND^{-1}}_{=-A^T} H \underbrace{D^{-1}N}_{=-A} \tag{4.14}$$

Hence:

$$\begin{aligned} -\frac{1}{2}u_p^T P u_p &= \frac{1}{2}u_p^T(A^T HA - H)u_p \\ P_1 &= V_1 \end{aligned} \tag{4.15}$$

We proceed to analyze the term P_2. Using the fact that $u^* = -H^{-1}c$, we have that:

$$\begin{aligned} &P_2 = \\ &\quad = -\frac{1}{2}u_p^T(HD^{-1}(D-N)D^{-1}H)H^{-1}c \\ &\quad = -\frac{1}{2}u_p^T(HD^{-1}(D-N)b) \end{aligned} \tag{4.16}$$

where we used the fact that $D^{-1}c = b$. Now observe that:

$$A^T H b = (D+N)b = (N+D)D^{-1}c \\ = (-A^T + I)c \tag{4.17}$$

Which by substitution in equation 4.16 yields:

$$\begin{aligned} P_2 &= \\ &= -\frac{1}{2}u_p^T(HD^{-1}(D-N)b) \\ &= \frac{1}{2}u_p^T(A^T H b + (A^T - I)c) = V_2 \end{aligned} \tag{4.18}$$

By which we can also infer that since $V_2 = V_3^T$ and $P_2 = P_3^T$ we must have that $P_3 = V_3$.

Lastly, we analyze the quadratic term (in u^*) P_4:

$$\begin{aligned} P_4 &= -\frac{1}{2}u^{*T}Pu^* \\ &= -\frac{1}{2}(c^T H^{-1} H D^{-1}(D-N)D^{-1} H H^{-1} c) \\ &= -\frac{1}{2}(c^T D^{-1}(D-N)D^{-1}c) \\ &= -\frac{1}{2}(c^T D^{-1} D D^{-1} c - c^T D^{-1} N D^{-1} c) \\ &= -\frac{1}{2}(c^T D^{-1} c - c^T D^{-1} N D^{-1} c) \\ &= -\frac{1}{2}(c^T D^{-1} c - c^T D^{-1}(H-D)D^{-1} c) \\ &= -\frac{1}{2}(2c^T D^{-1} c + c^T D^{-1} H D^{-1} c) \\ &= c^T b + \frac{1}{2} b^T H b = V_4 \end{aligned} \tag{4.19}$$

Again using the fact that $D^{-1}c = b$.

In conclusion this shows that:

$$V(u^{p+1}) - V(u_p) = -\frac{1}{2}(u_p - u^*)^T P(u_p - u^*) \prec 0 \tag{4.20}$$

which thereby ends the proof ■

References

[1] A. Hansen. *Mathematical Tripos Part II: Michaelmas Term 2020 Numerical Analysis – Lecture 17*. [Accessed: 2024-06-06]. URL: https://www.damtp.cam.ac.uk/research/afha/lectures/Part_II_NumAn/b17.pdf.

[2] Björn Johansson. "On Distributed Optimization in Networked Systems". QC 20100813. Doctoral thesis. Stockholm: KTH, School of Electrical Engineering (EES), Automatic Control, 2008, pp. 18, 55. ISBN: 978-91-7415-190-9. URL: https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-9801.