Delft University of Technology

Networked and distributed control systems

SC42101

Assignment 3

Authors:

Mihaly Fey (5476747)

June 20, 2025

1 Problem 1

Prove that $(z-x^*)^T(-\nabla f(x^*) \leq 0,\ \forall z \in \mathcal{X} \iff f(x) \geq f(x^*),\ \forall x \in \mathcal{X}$. $f(x)$ is convex. For convex functions Equation 1 holds.

$$f(x) \geq f(x_0) + (x-x_0)^T \nabla f(x_0) \tag{1}$$

To prove the implication:

$$(z-x^*)^T(\nabla f(x^*) \geq 0 \tag{2}$$
$$(z-x^*)^T(\nabla f(x^*) + f(x^*) \geq f(x^*) \tag{3}$$
$$f(z) \geq (z-x^*)^T(\nabla f(x^*) + f(x^*) \geq f(x^*) \tag{4}$$
$$f(z) \geq f(x^*), \forall z \in \mathcal{X} \tag{5}$$

To prove the contrapositive, let $y \in \mathcal{X}$ be chosen such, that $f(y) < f(x^*)$

$$f(y) < f(x^*) \tag{6}$$
$$f(x^*) + (y-x^*)^T(\nabla f(x^*) \leq f(y) \tag{7}$$
$$f(x^*) + (y-x^*)^T(\nabla f(x^*) \leq f(y) < f(x^*) \tag{8}$$
$$f(x^*) + (y-x^*)^T(\nabla f(x^*) < f(x^*) \tag{9}$$
$$(y-x^*)^T(\nabla f(x^*) < 0 \tag{10}$$
$$\exists x \in \mathcal{X}, s.t.(z-x^*)^T(-\nabla f(x^*) > 0 \tag{11}$$

$$(z-x^*)^T(-\nabla f(x^*) \leq 0,\ \forall z \in \mathcal{X} \iff f(x) \geq f(x^*),\ \forall x \in \mathcal{X} \tag{12}$$

2 Problem 2

The update step is the following:

$$\begin{bmatrix} u_1^{p+1} \\ u_2^{p+1} \end{bmatrix} = w_1 \begin{bmatrix} u_1^{p+}(u_2^p) \\ u_2^p \end{bmatrix} + w_2 \begin{bmatrix} u_1^p \\ u_2^{p+}(u_1^p) \end{bmatrix} \tag{13}$$

$$u_1^{p+}(u_2^p) = -H_{11}^{-1}(H_{12}u_2^p + c_1) \quad u_2^{p+}(u_1^p) = -H_{22}^{-1}(H_{21}u_2^p + c_2) \tag{14}$$

$$\begin{bmatrix} u_1^{p+1} \\ u_2^{p+1} \end{bmatrix} = \begin{bmatrix} -w_1 H_{11}^{-1} H_{12} u_2^p \\ w_1 u_2^p \end{bmatrix} + \begin{bmatrix} w_2 u_1^p \\ -w_2 H_{22}^{-1} H_{21} u_1^p \end{bmatrix} + \begin{bmatrix} -w_1 H_{11}^{-1} c_1 \\ -w_2 H_{22}^{-1} c_2 \end{bmatrix} \tag{15}$$

$$\begin{bmatrix} u_1^{p+1} \\ u_2^{p+1} \end{bmatrix} = \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} \begin{bmatrix} u_1^p \\ u_2^p \end{bmatrix} + \begin{bmatrix} -w_1 H_{11}^{-1} c_1 \\ -w_2 H_{22}^{-1} c_2 \end{bmatrix} = Au^p + b \tag{16}$$

It is easy to see that if either $w_1 = 0$ or $w_2 = 0$ the method does not converge, since one set of variables does not get updated, so in the followings, it is assumed, that $0 < w_1, w_2, < 1$.

To get the eigenvalues of A the equation $\det(A - \lambda I) = 0$ has to be solved. The determinant can be calculated with the help of the determinant formula of the Schur-complement $\det(M) = \det(A)\det(D - CA^{-1}B)$. If $\lambda = w_1$ or $\lambda = w_2$ the respective sub-matrices would not be invertable, so the other Schur-complement can be used, at which point it becomes clear, that both are eigenvalues of A. In the following it is assumed that $\lambda \neq w_1, w_2$. Furthermore $S(x)$ denotes the Schur-complement of H, with the respective submatrix.

$$\det(A - \lambda I) = \det((w_2 - \lambda)I)\det((w_1 - \lambda)I - \frac{w_1 w_2}{(w_2 - \lambda)} H_{22}^{-1} H_{21} H_{11}^{-1} H_{12}) = 0 \tag{17}$$

$$(w_2 - \lambda)(w_1 - \lambda)\det(I - \alpha M) = 0 \quad M = H_{22}^{-1} H_{21} H_{11}^{-1} H_{12} \ \alpha = \frac{w_1 w_2}{(\lambda - w_1)(\lambda - w_2)} \tag{18}$$

From the above expression it can be seen that $\frac{1}{\alpha}$ are the eigenvalues of M. Since $H > 0$, $S(H_{11}) > 0$, $H_{22}, H_{22}^{-1} > 0$, therefore $H_{22}^{-1} S(H_{11}) > 0$.

$$I - M = H_{22}^{-1} S(H_{11}) > 0, \quad \lambda(I - M) > 0 \tag{19}$$

$$\lambda(I - M) > 0 \implies \lambda(M - I) < 0 \implies \lambda(M - I + I) < 1 \implies \lambda(M) < 1 \implies \frac{1}{\alpha} < 1 \tag{20}$$

$$\frac{(\lambda - w_1)(\lambda - w_2)}{w_1 w_2} < 1 \tag{21}$$
$$(\lambda - w_1)(\lambda - w_2) < w_1 w_2 \tag{22}$$
$$\lambda^2 - (w_1 + w_2)\lambda + w_1 w_2 < w_1 w_2 \tag{23}$$
$$\lambda^2 - \lambda < 0 \tag{24}$$
$$0 < \lambda < 1 \tag{25}$$

Given that the iteration converges the only invariant point in $\mathcal{X}$ is the converged point, and all other starting points converge to this one point. As such if an invariant point is found that will be solution of the optimization problem. Assuming this point $u^* = -H^{-1}c$. To prove this is an invariant the follwing equaiton has to be true:

$$A(-H^{-1}c) + b = -H^{-1}c \tag{26}$$
$$HAH^{-1}c - Hb = c \tag{27}$$

Since both H_{11} and H_{22} are invertible both Schur-complements can be created. This means that the matrix inversion lemma can be applied to the Schur-inversion formula, resulting in the Equation 28.

$$H^{-1} = \begin{bmatrix} S(H_{22})^{-1} & -S(H_{22})^{-1}H_{12}H_{22}^{-1} \\ -S(H_{11})^{-1}H_{21}H_{11}^{-1} & S(H_{11})^{-1} \end{bmatrix} \tag{28}$$

$$HAH^{-1}c = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1}H_{12} \\ -w_2 H_{22}^{-1}H_{21} & w_1 I \end{bmatrix} \begin{bmatrix} S(H_{22})^{-1} & -S(H_{22})^{-1}H_{12}H_{22}^{-1} \\ -S(H_{11})^{-1}H_{21}H_{11}^{-1} & S(H_{11})^{-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{29}$$

$$HAH^{-1}c = \begin{bmatrix} w_2 S(H_{22}) & 0 \\ 0 & w_1 S(H_{11}) \end{bmatrix} \begin{bmatrix} S(H_{22})^{-1}c_1 - S(H_{22})^{-1}H_{12}H_{22}^{-1}c_2 \\ S(H_{11})^{-1}c_2 - S(H_{11})^{-1}H_{21}H_{11}^{-1}c_1 \end{bmatrix} \tag{30}$$

$$HAH^{-1}c - Hb = \begin{bmatrix} w_2 I c_1 - w_2 H_{12}H_{22}^{-1}c_2 \\ w_1 I c_2 - w_1 H_{21}H_{11}^{-1}c_1 \end{bmatrix} + \begin{bmatrix} w_1 I c_1 + w_2 H_{21}H_{22}^{-1}c_2 \\ w_2 I c_2 + w_1 H_{21}H_{11}^{-1}c_1 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{31}$$

3 Problem 3

Assuming:

$$D = \begin{bmatrix} w_1^{-1}H_{11} & 0 \\ 0 & w_2^{-1}H_{22} \end{bmatrix} \; N = \begin{bmatrix} -w_1^{-1}w_2 H_{11} & H_{12} \\ H_{21} & -w_1 w_2^{-1}H_{22} \end{bmatrix} \; \tilde{H} = \begin{bmatrix} H_{11} & -H_{12} \\ -H_{21} & H_{22} \end{bmatrix} \; P = HD^{-1}\tilde{H}D^{-1}H \tag{32}$$

Expressing the cost for the optimal value first, the coordinate change for u to $(u - u^*)$ is the following:

$$u^* = -H^{-1}cV(u^*) = \frac{1}{2}u^* H u^* + c^T u^* + d = \frac{1}{2} = c^T H^{-1} H H^{-1} c - c^T H^{-1} c + d \tag{33}$$

$$V(u^*) = -\frac{1}{2}c^T H^{-1} c \tag{34}$$

$$V(u^p) = \frac{1}{2}(u^* + (u^p - u^*))^T H (u^* + (u^p - u^*)) + c^T(u^* + (u^p - u^*)) + d \tag{35}$$

$$V(u^p) = \frac{1}{2}[u^{*T}Hu^* + (u^p - u^*)^T H u^* + u^{*T}H(u^p - u^*) + (u^p - u^*)^T H (u^p - u^*)] + c^T u^* + c^T(u^p - u^*) + d \tag{36}$$

$$V(u^p) = V(u*) + \frac{1}{2}[(u^p - u^*)^T H u^* + u^{*T}H(u^p - u^*) + (u^p - u^*)^T H (u^p - u^*)] + c^T(u^p - u^*) \tag{37}$$

V is scalar, so any term can be freely transposed, furthermore $H^T = H$, and $(c^T + u^{*T}H) = c^T - c^T = 0$.

$$V(u^p) = V(u^*) + \frac{1}{2}(u^p - u^*)^T H(u^p - u^*) + (c^T + u^{*T}H)(u^p - u^*) = V(u^*) + \frac{1}{2}(u^p - u^*)^T H(u^p - u^*) \quad (38)$$

The coordinate change for the update equation is the following:

$$u^{p+1} = Au^p + b = A(u^p - u^*) + Au^* + b = A(u^p - u^*) + u^* \quad (39)$$

$$u^{p+1} - u^* = A(u^p - u^*) \quad (40)$$

$$u^p - u^* = v^p \quad (41)$$

Putting all of this together:

$$\begin{gathered} \underset{v}{\text{minimize}} \quad V(v) = \frac{1}{2}v^T Hv \\ v^{p+1} = Av^p, \quad v^p = [v_1^{pT} \ v_2^{pT}]^T \\ A = \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} \end{gathered} \quad (42)$$

Substituting the $v_{p+1} = Av^p$ into the cost:

$$\frac{1}{2}v^{p+1T} H v^{p+1} = \frac{1}{2}v^{pT} H_{new} v^p = \frac{1}{2}v^{pT} A^T HAv^p \quad (43)$$

Then:

$$V(u^{p+1}) - V(u^p) = V(v^{p+1}) - V(v^p) = \frac{1}{2}v^{pT}(A^T HA - H)v^p \quad (44)$$

To prove: $H_{new} - H = -\frac{1}{2}P > 0$, also can be written as $H_{new}H^{-1} - I = -\frac{1}{2}PH^{-1} > 0$. All positive definite matrices are symmetric, and multiplication of symmetric square matrices is commutative. , therefore $A^T HA = A^T AH$. What is then left to prove is that $0.5(A^T A - I) = -0.5HD^{-1}\tilde{H}D^{-1}$

It can be noticed that $A = -D^{-1}N$, $H = D + N$, $\tilde{H} = D - N$

$$0.5(A^T A - I) = -0.5(D + N)D^{-1}(D - N)D^{-1} \quad (45)$$

$$ND^{-1}D^{-1}N - I = -I + D^{-1}NND^{-1} \quad (46)$$

$$ND^{-1}D^{-1}N - I = ND^{-1}D^{-1}N - I \quad (47)$$

Therefore $V(u^{p+1}) - V(u^p) = -\frac{1}{2}v^{pT}Pv^p$. To prove the decrease, it is enough to prove $-\frac{1}{2}P < 0$ as per the definition of a negative definite matrix the cost function will be decreasing. Since $H > 0$, $H_{11}, H_{22} > 0$, therefore $D > 0$. Then $D^{-1} > 0$. As per the problem, $\tilde{H} > 0$ and the product of positive definite matrices will be positive definite as well, so $P > 0$. Therefore $-\frac{1}{2}v^{pT}Pv^p < 0$.

4 Problem 4

The new optimization procedure, which optimizes the groups of variables separately, does not always show a decrease in the cost function. Taking the example shown in the assignment, Equation 48. Shows an increase in the cost function. Since the the problem is 3 dimensional, each single optimization can be performed on a scalar. The three cost functions to be minimized are Equation 49. The new u vector these optimizations produce are Equation 50. The cost function for this new vector is 6.75, while for the original u it is 4.0, which means the method does not converge. The values were calculated with the script `assignment_3.py`

$$H = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad c = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad u^p = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad (48)$$

$$\underset{u_1}{min} \ 2u_1^2 + u_1 \quad \underset{u_2}{min} \ u_2^2 + 3u_2 \quad \underset{u_3}{min} \ 2u_3^2 + 2u_3 \quad (49)$$

$$u^{p+1} = \begin{bmatrix} -0.5 \\ -3 \\ -1 \end{bmatrix} \quad (50)$$