Assignment 3

SC42101: Networked and Distributed Control Systems

Richard Adam [redacted]

Delft University of Technology

TU Delft

Introduction

Question 1: Proving a Condition of Optimality

Begin with the following convex optimization problem:

$$\min_x \quad f(x)$$
$$\text{subject to} \quad x \in \mathcal{X}$$

where

$$f(x) = \frac{1}{2}x^T H x + c^T x \tag{1}$$

where $H \succeq 0$ is some positive semi-definite matrix and $\mathcal{X}$ is a convex set. The goal is to prove that the term:

$$(z - x^*)(-\nabla f(x^*)) \leq 0, \quad \forall z \in \mathcal{X} \tag{2}$$

is both necessary and sufficient to prove that x^* is an optimal solution.

Proving Necessity

To prove necessity one must start with the definition of the optimal x, denoted as x^*, which is that

$$f(x^*) \leq f(z) \quad \forall\, z \in \mathcal{X} \tag{3}$$

we can then use the property of convexity, which is that:

$$ax + (1-a)y \in \mathcal{X}, \quad \text{for } a \in [0,1] \tag{4}$$

For any x or y in the convex set. Setting $x = z$ and $y = x^*$, one can create a new variable x_a and rearrange the variables to define:

$$x_a = x^* + a(x^* - z) \in \mathcal{X}, \quad \text{for } a \in [0,1] \tag{5}$$

By the definition of optimality, we know that $f(x^*) \leq f(x_a)$, so $f(x_a) - f(x^*) \geq 0$. Plugging x_a into Equation 1, one obtains:

$$\begin{aligned}
f(x_a) &= \frac{1}{2}(x^* + a(z - x^*))^T H(x^* + a(z - x^*)) + c^T(x^* + a(z - x^*)) \\
&= \frac{1}{2}(x^* + a(z - x^*))^T[Hx^* + aH(z - x^*)] + c^T x^* + ac^T(z - x^*)) \\
&= \frac{1}{2}x^{*T}Hx^* + \frac{1}{2}a(z - x^*)^T Hx^* + \frac{1}{2}ax^{*T}H(z - x^*) + \frac{1}{2}a^2(z - x^*)^T H(z - x^*) + c^T x^* + ac^T(z - x^*))
\end{aligned}$$

Since H is a positive semi-definite matrix, by definition it is symmetric, so this expression can be further simplified to:

$$\begin{aligned}
f(x_a) &= \frac{1}{2}x^{*T}Hx^* + ax^{*T}H(z - x^*) + \frac{1}{2}a^2(z - x^*)^T H(z - x^*) + c^T x^* + ac^T(z - x^*)) \\
&= (\frac{1}{2}x^* Hx^* + c^T x^*) + ax^{*T}H(z - x^*) + \frac{1}{2}a^2(z - x^*)^T H(z - x^*) + ac^T(z - x^*)) \\
&= f(x^*) + ax^{*T}H(z - x^*) + \frac{1}{2}a^2(z - x^*)^T H(z - x^*) + ac^T(z - x^*))
\end{aligned}$$

now, one can express the inequality:

$$f(x_a) - f(x^*) = ax^{*T}H(z - x^*) + \frac{1}{2}a^2(z - x^*)^T H(z - x^*) + ac^T(z - x^*)) \geq 0$$

$$a(x^{*T}H + c^T)(z - x^*) + \frac{1}{2}a^2(z - x^*)^T H(z - x^*) \geq 0$$

Divide by a

$$(x^{*T}H + c^T)(z - x^*) + \frac{1}{2}a(z - x^*)^T H(z - x^*) \geq 0$$

Taking the gradient of $f(x)$, one gets the equation $\nabla f(x) = Hx + c$, so $\nabla f(x^*) = Hx^* + c$, meaning $(x^{*T}H + c^T) = \nabla f(x^*)^T$. With this one can rewrite the inequality as:

$$(\nabla f(x^*)^T(z - x^*) + \frac{1}{2}a(z - x^*)^T H(z - x^*) \geq 0$$

To examine a point very close to x^*, one can take the limit as a$\to$0.

$$\lim_{a\to 0}[f(x_a) - f(x^*)] = \lim_{a\to 0} \nabla[f(x^*)^T(z - x^*) + \frac{1}{2}a(z - x^*)^T H(z - x^*)] = \nabla f(x^*)^T(z - x^*) \geq 0$$

This can be rewritten as $(z - x^*)^T\nabla f(x^*) \geq 0$, which can then be expressed as:

$$(z - x^*)^T(-\nabla f(x^*)) \leq 0$$

This proves that, in order for the solution x^* to be optimal, the inequality $(z - x^*)^T(-\nabla f(x^*)) \leq 0, \quad \forall\, z \in \mathcal{X}$ is necessarily true.

Proving Sufficiency

To prove sufficiency, one can start with the equation:

$$(z - x^*)^T(-\nabla f(x^*)) \leq 0$$

Recalling the affine lower bound property of convex functions:

$$f(z) \geq f(x^*) + \nabla f(x^*)(z - x^*)^T$$

it is already established that $(z - x^*)^T(-\nabla f(x^*)) \leq 0$, which can be converted to $\nabla f(x^*)(z - x^*)^T \geq 0$. Using this inequality, one knows that:

$$f(z) \geq f(x^*) + \nabla f(x^*)(z - x^*)^T \geq f(x^*) + 0$$

Thus, using the transitivity of inequalities:

$$f(z) \geq f(x^*) \quad \forall\, z \in \mathcal{X}$$

This is the definition of optimality, thus the fact that $(z - x^*)^T(-\nabla f(x^*)) \leq 0 \quad \forall z \in \mathcal{X}$ is sufficient to say that the solution x^* is optimal. Since both necessity and sufficiency have been proven, one say that x^* is an optimal solution if and only if $(z - x^*)^T(-\nabla f(x^*)) \leq 0 \quad \forall z \in \mathcal{X}$. Looking back on this graphically, this condition can be thought of as saying that for all possible solutions z, moving from x^* to z is moving up the gradient of the objective function $f(x)$, meaning the value $f(z)$ will either be higher or equal to $f(x^*)$.

Question 2: Optimization via Partitioning

This question consideres partitioning the optimization problem: $V(u) = \frac{1}{2}u^T Hu + c^T u + d$ into the following problem:

$$V(u_1, u_2) = \frac{1}{2}[u_1^T \quad u_2^T] \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + [c_1^T \quad c_2^T] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + d \tag{6}$$

The problem will be iteratively optimized by first optimizing over u_1 while keeping u_2 constant to find u_1^*, then optimizing over u_2 while keeping u_1 constant to find u_2^*, the defining the final variable as a convex combination of the two optimized variables.

Part A: Finding Iterative Expression

The first step in finding the expression for u^{p+1} from u^p is to optimize over u_1 while keeping u_2 constant, in other words:

$$\min_{u_1} \frac{1}{2}[u_1^T \quad u_2^T] \begin{bmatrix} H11 & H12 \\ H21 & H22 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + [c_1^T \quad c_2^T] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + d$$

The matrix can be expanded to:

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

Using the fact that H is symmetric, which means $H_{12} = H_{21}^T$, one can simplify the expression to:

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

One can minimize the function with respect to u_1 by solving for $\frac{\partial V}{\partial u_1} = 0$, or:

$$\frac{\partial V}{\partial u_1} = H_{11} u_1 + H_{12} u_1 + c_1 = 0$$

Solving this for u_1^*, one obtains:

$$u_{1,\text{step 1}}^* = -H_{11}^{-1}(H_{12} u_2 + c_1),$$

making the entire expression for step 1:

$$u_{\text{step 1}}^* = \begin{bmatrix} -H_{11}^{-1}(H_{12} u_2 + c_1) \\ u_2 \end{bmatrix}$$

The next step is to optimize over u_2 while keeping u_1 constant, in other words:

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

One can minimize the function with respect to u_1 by solving for $\frac{\partial V}{\partial u_2} = 0$, or:

$$\frac{\partial V}{\partial u_2} = u_1^T H_{12} + H_{22} u_2 + c_2 = H_{21} u_1 + H_{22} u_2 + c_2 = 0$$

solving this for u_2 results in the expression:

$$u^*_{2,\text{step 2}} = -H_{22}^{-1}(H_{21}u_1 + c_2)$$

Which means the final expression for step 2 is:

$$u^*_{\text{step 2}} = \begin{bmatrix} u_1 \\ -H_{22}^{-1}(H_{21}u_1 + c_2) \end{bmatrix}$$

The final step is to combine the two in a convex combination, as follows:

$$u^{p+1} = w_1 u^*_{\text{step 1}} + w_2 u^*_{\text{step 2}} = w_1 \begin{bmatrix} -H_{11}^{-1}(H_{12}u_2 + c_1) \\ u_2 \end{bmatrix} + w_2 \begin{bmatrix} u_1 \\ -H_{22}^{-1}(H_{21}u_1 + c_2) \end{bmatrix}$$

Where $w_1 + w_2 = 1$. Expanding the expression, one can obtain the two expressions:

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

This can be expressed by the linear system:

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

Part B: Proving Convergence

In order to prove that the convergence of the expression $x^{p+1} = Ax^p + B$, one must only confirm that the spectral radius of A is less than 1. This is broken into two parts, first ensuring that all eigenvalues all eigenvalues are less than 1, and second ensuring that the eigenvalues are positive.

Proving that Eigenvalues are Less Than 1

To find an upper bound for the eigenvalues, one can start with the simple quadratic relation:

$$\begin{aligned} Q(\nu) = [\nu_1^T \ \nu_2^T]\bar{H} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} &= [\nu_1^T \ \nu_2^T] \begin{bmatrix} H_{11} & -H_{12} \\ -H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \\ &= \nu_1^T H_{11} \nu_1 + \nu_2^T H_{22} \nu_2 - \nu_1^T H_{12} \nu_2 - \nu_2^T H_{21} \nu_1 \\ &= \nu_1^T H_{11} \nu_1 + \nu_2^T H_{22} \nu_2 - \nu_1^T H_{12} \nu_2 - \nu_2^T H_{12}^T \nu_1 \\ &= \nu_1^T H_{11} \nu_1 + \nu_2^T H_{22} \nu_2 - 2\nu_1^T H_{12} \nu_2 \end{aligned}$$

Because $\bar{H} \succ 0$, by definition $Q(\nu) > 0$ for all ν values. Next, one can express the general eigenvalue-eigenvector relationship $A\nu = \lambda\nu$ as:

$$\begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \lambda \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix}$$

Expanding the equation, one can write this as:

$$\begin{aligned} w_2 \nu_1 - w_1 H_{11}^{-1} H_{12} \nu_2 &= \lambda \nu_1 \\ w_1 \nu_2 - w_2 H_{22}^{-1} H_{12} \nu_1 &= \lambda \nu_2 \end{aligned}$$

Isolating each variable results in the equations:

$$(w_2-\lambda)\nu_1 = w_1 H_{11}^{-1} H_{12}\nu_2 \tag{7}$$

$$(w_1-\lambda)\nu_2 = w_2 H_{22}^{-1} H_{12}\nu_1 \tag{8}$$

Equation 7 can be rearranged as:

$$H_{12}\nu_2 = \frac{w_2-\lambda}{w_1} H_{11}v_1$$

Premultiply by ν_1^T:

$$\nu_1^T H_{12}\nu_2 = \frac{w_2-\lambda}{w_1}\nu_1^T H_{11}v_1$$

Similarly, Equation 8 can be written as:

$$\nu_2 = \frac{w_2}{w_1-\lambda} H_{22}^{-1} H_{21}v_1$$

Premultiply by $\nu_2^T H_{22}$:

$$\begin{aligned}
\nu_2^T H_{22}\nu_2 &= \frac{w_2}{w_1-\lambda}\nu_2^T H_{22} H_{22}^{-1} H_{21}v_1 \\
\nu_2^T H_{22}\nu_2 &= \frac{w_2}{w_1-\lambda}\nu_2^T H_{21}v_1 \\
\nu_2^T H_{22}\nu_2 &= \frac{w_2}{w_1-\lambda}\nu_1^T H_{12}v_2 \\
\nu_2^T H_{22}\nu_2 &= (\frac{w_2}{w_1-\lambda})(\frac{w_2-\lambda}{w_1})\nu_1^T H_{11}v_1
\end{aligned}$$

These two expressions can be plugged into the formula for the quadratic function $Q(\nu)$:

$$\begin{aligned}
Q(\nu) &= \nu_1^T H_{11}\nu_1 + \nu_2^T H_{22}\nu_2 - 2\nu_1^T H_{12}\nu_2 \\
&= \nu_1^T H_{11}\nu_1 + (\frac{w_2}{w_1-\lambda})(\frac{w_2-\lambda}{w_1})\nu_1^T H_{11}v_1 - 2\frac{w_2-\lambda}{w_1}\nu_1^T H_{11}v_1 \\
&= (1 + (\frac{w_2}{w_1-\lambda})(\frac{w_2-\lambda}{w_1}) - 2\frac{w_2-\lambda}{w_1})\nu_1^T H_{11}v_1
\end{aligned}$$

Solvingthe expression within the parenthesis, one obtains the equation:

$$\begin{aligned}
Q(\nu) &= (\frac{w_1^2 - w_1\lambda + 2w_1w_2 - 2w_2\lambda - 2w_1\lambda + 2\lambda^2 + w_2^2 - w_2\lambda}{w_1(w_1-\lambda)})\nu_1^T H_{11}v_1 \\
&= (\frac{(w_1^2 + w_2^2 + 2w_1w_2) - (w_1+w_2)\lambda - 2(w_1+w_2)\lambda + 2\lambda^2}{w_1(w_1-\lambda)})\nu_1^T H_{11}v_1 \\
&= (\frac{(w_1+w_2)^2 - 3\lambda + 2\lambda^2}{w_1(w_1-\lambda)})\nu_1^T H_{11}v_1 \\
&= (\frac{1 - 3\lambda + 2\lambda^2}{w_1(w_1-\lambda)})\nu_1^T H_{11}v_1
\end{aligned}$$

One knows that, since H is a positive definite matrix, $Q(\nu) > 0$ for all $\nu \neq 0$ Since H_{11} is a positive definite matrix, the term $\nu_1^T H_{11}v_1 > 0$ by definition, which means the coefficient term $\frac{1-3\lambda+2\lambda^2}{w_1(w_1-\lambda)} > 0$. Solving the quadratic equation results in the roots of 1 and 0.5, so $\lambda < 1$ for the given A matrix.

Proving that Eigenvalues are Greater Than 0
Another way to express the eigenvalue definition is using the form:

$$\det(A - \lambda I) = 0$$

Solving $A - \lambda I$ results in the matrix:

$$A - \lambda I = \begin{bmatrix} (w_2 - \lambda)I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & (w_1 - \lambda)I \end{bmatrix}$$

In order to solve the determinant, one can use the Schur Complement decomposition of the determinant of a 2x2 block matrix, which is:

$$\det(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}) = \det(A_{11}) \det(A_{22} - A_{21} A_{11}^{-1} A_{12})$$

Applying this to the $A - \lambda I$ matrix, one obtains:

$$\begin{aligned} &\det(\begin{bmatrix} (w_2 - \lambda)I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & (w_1 - \lambda)I \end{bmatrix}) \\ &= \det((w_2 - \lambda)I) \det((w_1 - \lambda)I + (w_2 H_{22}^{-1} H_{21})((w_2 - \lambda)I)^{-1}(-w_1 H_{11}^{-1} H_{12})) \\ &= \det((w_2 - \lambda)I) \det((w_1 - \lambda)I + \frac{w_2 w_1}{w_2 - \lambda}(H_{22}^{-1} H_{12}^T H_{11}^{-1} H_{12})) \end{aligned}$$

Taking the first determinant as 0 is simple, as it results in eigenvalues of w_2, which is between zero and one by definition. Taking the second determinant requires us to know that, since H_{11} and H_{22} are positive definite matrices, their inverse is also positive definite. This means that $H_{12}^T H_{11}^{-1} H_{12}$ is also positive definite, by the theorem that for any positive definite H, $R^T H R \succ 0$ for some non-singular R matrix. In addition, the product of two symmetric positive definite matrices is positive definite, so the entire term $H_{22}^{-1} H_{12}^T H_{11}^{-1} H_{12}$ is positive definite. Combined with the fact that the identity matrix is positive definite, that means for the determinant to be zero, the terms $(w_2 - \lambda)$ and $\frac{w_2 w_1}{w_2 - \lambda}$ must have opposite signs. First considering the case where $\lambda > w_1$, that means that λ must be less than w_2 to ensure opposite signs, and when $\lambda > w_2$, that means that λ must be less than w_1. This means that the lambda is bounded on the upper and lower end by the possible values of w_1 and w_2, so $\lambda > 0$. Combining this with the fact that $\lambda < 1$ from the previous section, one can confirm that the solution will converge, since the spectral radius of A is less than 1.

Part C: Expression for Converged Optimum

With the knowledge that this method will converge, we can now confirm that it reaches the same solution as $u^* - -H^{-1}c$, which is obtained by setting the gradient of the function equal to zero and solving for u.

To confirm this, one can use the expectation that at the optimal solution, the next iteration will return the optimal solution again, such that $x^{p+1} = x^p = x^*$. Using the expressions for x_1^{p+1} and x_2^{p+1}, this is expressed as:

$$\begin{aligned} u_1^* &= -w_1 H_{11}^{-1} H_{12} u_2^* - w_1 H_{11}^{-1} c_1 + w_2 u_1^* \\ u_2^* &= w_1 u_2^* - w_2 H_{22}^{-1} H_{21} u_1^* - w_2 H_{22}^{-1} c_2 \end{aligned}$$

Rearranging the optimality condition as $-c = Hu^*$, the c components can be rewritten as:

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

Plugging these expression into the above equation, one obtains:

$$\begin{aligned} u_1^* &= -w_1 H_{11}^{-1} H_{12} u_2^* + w_1 H_{11}^{-1}(H_{11}u_1^* + H_{12}u_2^*) + w_2 u_1^* \\ u_2^* &= w_1 u_2^* - w_2 H_{22}^{-1} H_{21} u_1^* + w_2 H_{22}^{-1}(H_{21}u_1^* + H_{22}u_2^*) \end{aligned}$$

Which simplifies to:

$$\begin{aligned} u_1^* &= -w_1 H_{11}^{-1} H_{12} u_2^* + w_1 H_{11}^{-1} H_{11} u_1^* + w_1 H_{11}^{-1} H_{12} u_2^* + w_2 u_1^* = (w_1 + w_2) u_1^* \\ u_2^* &= w_1 u_2^* - w_2 H_{22}^{-1} H_{21} u_1^* + w_2 H_{22}^{-1} H_{21} u_1^* + w_2 H_{22}^{-1} H_{22} u_2^* = (w_1 + w_2) u_2^* \end{aligned}$$

Using the convex combination property $w_1 + w_2 = 1$, one finalizes the expressions as $u_1^* = u_1^*$ and $u_2^* = u_2^*$ respectively, confirming that this method converges to $u^* = -H^{-1}c$.

Question 3: Properties of Iteration

In this question, the goal is to find the size of decrease in the cost function defined in Equation 6 from p to p+1. Next, the goal is to prove that the cost function is monotonically decreasing, or:

$$V(u^{p+1}) < V(u^p)$$

Part A: Expressing Size of Decrease

The first step is to redefine the coordinate system centered around the optimal point u^*. In order to do this, one can define an intermediate variable $y^p = u^p - u^*$. Reversing this for u^p, one can rewrite the cost function as:

$$\begin{aligned} V(u) = V(y + u^*) &= \frac{1}{2}(y + u^*)^T H (y + u^*) + c^T (y + u^*) + d \\ &= \frac{1}{2}(y^T H y + u^{*T} H u^* + y^T H u^* + u^{*T} H y) + c^T y + c^T u^* + d \\ &\text{Using the fact that H is diagonal:} \quad y^T H u^* = u^{*T} H y \\ &= \frac{1}{2} y^T H y + \frac{1}{2} u^{*T} H u^* + u^{*T} H y + c^T y + c^T u^* + d \end{aligned}$$

Using the definition of optimality at $u^* = -H^{-1}c$, one can plug in $-c^T = -u^{*T}H$ to find:

$$\begin{aligned} V(u) = V(y + u^*) &= \frac{1}{2} y^T H y + \frac{1}{2} u^{*T} H u^* - c^T y + c^T y + c^T u^* + d \\ &= \frac{1}{2} y^T H y + \frac{1}{2} u^{*T} H u^* + u^{*T} H u^* + d \\ &= \frac{1}{2} y^T H y + V(u^*) \end{aligned}$$

The iteration difference can now be expressed in the new coordinate system as:

$$\begin{aligned} V(u^{p+1}) - V(u^p) &= V(y^{p+1} + u^*) - V(y^p + u^*) \\ = \frac{1}{2}y^{(p+1)T}Hy^{p+1} &+ V(u^*) - \frac{1}{2}y^{(p)T}Hy^p - V(u^*) \\ &= \frac{1}{2}y^{(p+1)T}Hy^{p+1} - \frac{1}{2}y^{(p)T}Hy^p \end{aligned}$$

One can find the expression for y^{p+1} in terms of y^p in the following way:

$$\begin{aligned} y^{p+1} &= u^{p+1} - u^* \\ y^{p+1} &= Au^p + b - u^* \\ \textbf{recall } u^p &= y^p + u^* : \\ y^{p+1} &= A(y^p + u^*) + b - u^* \\ y^{p+1} &= Ay^p + Au^* + b - u^* \end{aligned}$$

One can expand the b term by reversing the property $u^* = -H^{-1}c$ to $-c = Hu^*$. Expressing c_1 and c_2 individually, one can express this as $-c_1 = H_{11}u_1^* + H_{12}u_2^*$ and $-c_2 = H_{21}u_1^* + H_{22}u_2^*$. Plugging these into the equation for b obtains the expression:

$$b = \begin{bmatrix} -w_1H_{11}^{-1}c_1 \\ -w_2H_{22}^{-1}c_2 \end{bmatrix} = \begin{bmatrix} w_1H_{11}^{-1}H_{11}u_1^* + w_1H_{11}^{-1}H_{12}u_2^* \\ w_2H_{22}^{-1}H_{21}u_1^* + w_2H_{22}^{-1}H_{22}u_2^* \end{bmatrix} = \begin{bmatrix} w_1I & w_1H_{11}^{-1}H_{12} \\ w_2H_{22}^{-1}H_{21} & w_2I \end{bmatrix} u^*$$

Plugging this into the equation for y^{p+1}, one obtains:

$$\begin{aligned} y^{p+1} &= Ay^p + Au^* + b - u^* \\ = Ay^p + \begin{bmatrix} w_2I & -w_1H_{11}^{-1}H_{12} \\ -w_2H_{22}^{-1}H_{21} & w_1I \end{bmatrix} u^* &+ \begin{bmatrix} w_1I & w_1H_{11}^{-1}H_{12} \\ w_2H_{22}^{-1}H_{21} & w_2I \end{bmatrix} u^* - u^* \\ = Ay^p &+ \begin{bmatrix} (w_1+w_2)I & 0 \\ 0 & (w_1+w_2)I \end{bmatrix} u^* - u^* \\ = Ay^p &+ \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} u^* - u^* = Ay^p \end{aligned}$$

With the simple expression $y^{p+1} = Ay^p$, we can express the cost difference as:

$$V(u^{p+1}) - V(u^p) = \frac{1}{2}y^{(p+1)T}Hy^{p+1} - \frac{1}{2}y^{(p)T}Hy^p \tag{9}$$

$$= \frac{1}{2}(Ay^p)^T HAy^p - \frac{1}{2}y^{(p)T}Hy^p \tag{10}$$

$$= \frac{1}{2}y^{(p)T}A^T HAy^p - \frac{1}{2}y^{(p)T}Hy^p \tag{11}$$

$$= \frac{1}{2}y^{(p)T}(A^T HA - H)y^p \tag{12}$$

One can then express the function as:

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

Where:

$$P = HD^{-1}\tilde{H}D^{-1}H, \quad \tilde{H} = D - N, \quad H = D + N$$
$$D = \begin{pmatrix} w_1^{-1}H_{11} & 0 \\ 0 & w_2^{-1}H_{22} \end{pmatrix}, \quad N = \begin{pmatrix} -w_1^{-1}w_2H_{11} & H_{12} \\ H_{21} & -w_1w_2^{-1}H_{22} \end{pmatrix}$$

by finding $(A^THA - H) = -P$.

Part B: Proving Monotonic Decrease

Now that the cost function change is in the form:

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

Where:

$$P = HD^{-1}\tilde{H}D^{-1}H, \quad \tilde{H} = D - N, \quad H = D + N$$
$$D = \begin{pmatrix} w_1^{-1}H_{11} & 0 \\ 0 & w_2^{-1}H_{22} \end{pmatrix}, \quad N = \begin{pmatrix} -w_1^{-1}w_2H_{11} & H_{12} \\ H_{21} & -w_1w_2^{-1}H_{22} \end{pmatrix}$$

One can prove that the cost decreases monotonically by proving $P \succ 0$.

To begin, one can find the expression for $\tilde{H}$, which is:

$$\begin{aligned} \tilde{H} = D - N &= \begin{bmatrix} w_1^{-1}H_{11} & 0 \\ 0 & w_2^{-1}H_{22} \end{bmatrix} - \begin{bmatrix} -w_1^{-1}w_2H_{11} & H_{12} \\ H_{21} & -w_1w_2^{-1}H_{22} \end{bmatrix} \\ &= \begin{bmatrix} w_1^{-1}H_{11} + w_1^{-1}w_2H_{11} & -H_{12} \\ -H_{21} & w_2^{-1}H_{22} + w_1w_2^{-1}H_{22} \end{bmatrix} \end{aligned}$$

Simplifying this expression, one obtains:

$$\tilde{H} = \begin{bmatrix} \frac{1+w_2}{1-w_2}H_{11} & -H_{12} \\ -H_{21} & \frac{1+w_1}{1-w_1}H_{22} \end{bmatrix}$$

Since H is a positive definite matrix, one can use the property:

$$H = \begin{bmatrix} H_{11} & H_{11} \\ H_{21} & H_{22} \end{bmatrix} \succ 0, \quad \text{then} \quad \bar{H} = \begin{bmatrix} H_{11} & -H_{11} \\ -H_{21} & H_{22} \end{bmatrix} \succ 0$$

In order to know that the matrix $\tilde{H}$ would be positive definite without the $\frac{1+w_2}{1-w_2}$ and $\frac{1+w_1}{1-w_1}$ coefficients. One also knows that these coefficients are bounded to be between one and infinity since w_1 and w_2 are both bounded to be between 0 and 1. Since these coefficients are greater than 1 and they multiply the positive-definite diagonal block components of the matrix $\tilde{H}$, it is certain that $\tilde{H} \succ 0$.

Knowing that $\tilde{H} \succ 0$ allows one to use another property of positive definite matrices, which is that for the positive definite function $\tilde{H}$, the following expression will hold:

$$R^T\tilde{H}R \succ 0 \quad \text{for any nonsingular R matrix}$$

Since D is diagonal, D^{-1} will also be diagonal, so both H and D are symmetric matrices, meaning:

$$P = HD^{-1}\tilde{H}D^{-1}H = H^T(D^{-1T}\tilde{H}D^{-1})H$$

Using the previously stated property of positive definite matrices, one can now confirm that P is positive definite using the following chain:

$$\begin{aligned} \tilde{H} &\succ 0 \\ D^{-1T}\tilde{H}D^{-1} &\succ 0 \\ P = H^T(D^{-1T}\tilde{H}D^{-1})H &\succ 0 \end{aligned}$$

This means that P is positive definite, so the expression $\Delta V = -\frac{1}{2}(u^p - u^*)^T P(u^p - u^*)$ means the cost function $V(u)$ is monotonically decreasing, as by the definition of a positive definite matrix $(u^p - u^*)^T P(u^p - u^*) > 0$, so $-\frac{1}{2}(u^p - u^*)^T P(u^p - u^*) < 0 \quad \forall u^p \neq u^*$.

Question 4: Tri-Partitioned System Analysis

Going back to the cost function $V(u) = \frac{1}{2}u^T Hu + c^T u$. Instead of partitioning this into two parts, optimizing each, and taking a convex combination of the resulting optimal solutions, a newly proposed method is to partition into three parts, u_1, u_2, and u_3, then optimizing over each variable while keeping the other two constant, and using the outputs from each optimization as the starting point of the next step. In other words:

$$\begin{aligned} u_1^{p+1} &= \arg\min_{u_1} V(u_1, u_2^p, u_3^p) \\ u_2^{p+1} &= \arg\min_{u_2} V(u_1^p, u_2, u_3^p) \\ u_3^{p+1} &= \arg\min_{u_3} V(u_1^p, u_2^p, u_3) \end{aligned}$$

The same question as the double partition is posed: Will this guarantee monotonic decrease? In other words, is the statement $V(u_1^{p+1}, u_2^{p+1}, u_3^{p+1}) \leq V(u_1^p, u_2^p, u_3^p) \quad \forall u^p, u^{p+1}$ valid for all iterations? To examine this the following case has been generously provided:

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

Opimizing over u_1

First, one must check the value of $V(u^p)$, which can be expressed as:

$$V(u^p) = \frac{1}{2}u^{pT}Hu^p + c^T u^p = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 4$$

Now, to find u^p, one can first solve for u_1^p via solving:

$$x_1^{p+1} = \arg\min_{u_1} \frac{1}{2} \begin{bmatrix} u_1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \\ 1 \end{bmatrix}$$
$$= \frac{1}{2}(u_1^2 + u_1 + 2)$$

One can solve this expression by finding $\frac{\partial V}{\partial u_1} = 0$, or:

$$\frac{\partial V}{\partial u_1} = 2u_1 + 1 = 0, \quad u_1^{p+1} = -\frac{1}{2}$$

Optimizing over u_2

Next, optimizing over u_2 consists of solving:

$$x_2^{p+1} = \arg\min_{u_2} \frac{1}{2} \begin{bmatrix} 1 & u_2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ u_2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ u_2 \\ 1 \end{bmatrix}$$
$$= \frac{1}{2}u_2^2 + 3u_2 + 4$$

Again, one can minimize the function by finding $\frac{\partial V}{\partial u_2} = 0$, which can be seen in:

$$\frac{\partial V}{\partial u_2} = u_2 + 3 = 0, \quad u_2^{p+1} = -3$$

Optimizing over u_3

Lastly, optimizing over u_3:

$$x_3^{p+1} = \arg\min_{u_3} \frac{1}{2} \begin{bmatrix} 1 & 0 & u_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ u_3 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ u_3 \end{bmatrix}$$
$$= u_3^2 + 2u_3 + 1$$

$$\frac{\partial V}{\partial u_3} = 2u_3 + 2 = 0, \quad u_3^{p+1} = -1$$

Now that the expression $u^{p+1} = \begin{bmatrix} -\frac{1}{2} & -3 & -1 \end{bmatrix}^T$ is found, one can plug this into the cost equation to find:

$$V(u^p) = \frac{1}{2}u^{pT}Hu^p + c^T u^p = \frac{1}{2} \begin{bmatrix} -\frac{1}{2} & -3 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ -3 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ -3 \\ -1 \end{bmatrix} = 6.75$$

Comparing $V(u^p) = 4$ to $V(u^{p+1}) = 6.75$, one can clearly see that$V(u^{p+1}) \geq V(u^p)$, so for the tri-partitioned system with each variable being optimized independently, iterations do not guarantee monotonic decreases in the cost function.

Appendix A: Handwritten Notes and Math

Networked & Distributed Control Assignment 3

Problem 2

Considering the optimization problem

$$\underset{x}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad x \in \mathcal{X}$$

where $f(x) = \frac{1}{2}x^T H x + c^T x$, $H \succeq 0$ & $\mathcal{X}$ is a convex set

Show that x^* is an optimal solution iff

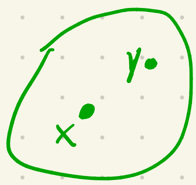
$$(z - x^*)^T(-\nabla f(x^*)) \leq 0 \;, \quad \forall z \in \mathcal{X}$$

all z within the possible set of x

Break Problem into 2 parts, "Necessity", & "Sufficiency"

① Necessary

by definition, we know $f(x^*) \geq f(z)$

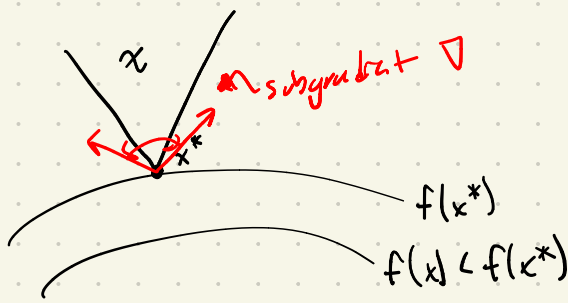


Important Note About Convex Sets:

we know that $ax + (1-a)y \in \mathcal{X}$ $\quad y + a(x-y) \in \mathcal{X}$

Consider a given point $x_a = x^* + a(x^* - z)$ for $a \in (0,1]$, $x_a \in \mathcal{X}$

again, by optimality we know $f(x^*) \leq f(x_a)$



by this we can say that:

$$f(x_a) - f(x^*) \geq 0$$

$$f(x_a) = \frac{1}{2}(x^* + a(z-x^*))^T H (x^* + a(z-x^*)) + c^T(x^* + a(z-x^*))$$

$$= \frac{1}{2}(x^{*T} + a(z-x^*)^T)[Hx^* + aH(z-x^*)] + c^T x^* + a c^T(z-x^*)$$

$$= \frac{1}{2}x^{*T} H x^* + \frac{1}{2}a(z-x^*)^T H x^* + \frac{1}{2}a x^{*T} H(z-x^*) + \frac{1}{2}a^2(z-x^*)^T H(z-x^*) +$$

H = PSD matrix, so it is symmetric by (some) definitions, hence

$$f(x_a) = \frac{1}{2}x^{*T} H x^* + a x^{*T} H(z-x^*) + \frac{1}{2}a^2(z-x^*)^T H(z-x^*) + c^T x^* + a c^T(z-x^*)$$

$$= \left(\frac{1}{2}x^{*T} H x^* + c^T x^*\right) + a x^{*T} H(z-x^*) + \frac{1}{2}a^2(z-x^*)^T H(z-x^*) + a c^T(z-x^*)$$

$$= f(x^*) + \alpha x^{*T} H(z-x^*) + \tfrac{1}{2}\alpha^2 (z-x^*)^T H(z-x^*) + \alpha c^T(z-x^*)$$

$$f(x_\alpha) - f(x^*) = \alpha x^{*T} H(z-x^*) + \tfrac{1}{2}\alpha^2 (z-x^*)^T H(z-x^*) + \alpha c^T(z-x^*)$$

$$= \alpha\left(x^{*T}H + c^T\right)(z-x^*) + \tfrac{1}{2}\alpha^2 (z-x^*)^T H(z-x^*)$$

We know that $\nabla f(x) = Hx + c$, so $\nabla f(x^*) = Hx^* + c = \nabla f(x^*)^T = x^{*T}H + c^T$

$$f(x_\alpha) - f(x^*) = \alpha \nabla f(x)^T (z-x^*) + \tfrac{1}{2}\alpha^2 (z-x^*)^T H(z-x^*) \geq 0$$

Divide by α:

$$\nabla f(x)^T (z-x^*) + \tfrac{1}{2}\alpha (z-x^*)^T H(z-x^*) \geq 0$$

Examining a point very close to x^*, i.e. $\alpha \to 0$

$$\lim_{\alpha\to 0}\left(f(x_\alpha) - f(x^*)\right) = \nabla f(x)^T (z-x^*) \geq 0$$

$$(z-x^*)^T \nabla f(x) \geq 0, \text{ so } \boxed{(z-x^*)^T(-\nabla f(x)) \leq 0}$$

Necessity established

Sufficiency: Starting @ $(z-x^*)^T(-\nabla f(x)) \leq 0$, getting to optimality

$\nabla f(x)(z-x^*)^T \geq 0$

first-order condition for convexity: $f(z) \geq f(x^*) + \nabla f(x^*)(z-x^*)^T$

– Affine lower bound

$$f(x^*) + \nabla f(x^*)(z-x^*)^T \geq f(x^*) + 0 \quad (\text{bcs } \nabla f(x)(z-x^*)^T \geq 0)$$

$$f(z) \geq f(x^*) + \nabla f(x^*)(z-x^*)^T \geq f(x^*) + 0$$

So

$f(z) \geq f(x^*) + 0$, $f(z) \geq f(x^*)\ \forall z \in \mathcal{X}$ ✓

Q2

Consider the positive-definite Quadratic fctn & partition into two variables:

$$V(u) = \frac{1}{2} u^T H u + c^T u + d$$

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

Optimize this in the following fashion:

1. optimize over u_1 while holding u_2 fixed to find u_1^*
2. optimize over u_2 while holding u_1 fixed to find u_2^*
3. make u^* convex combination of u_1^* & u_2^*

a.) Start From initial point u_1^p, u_2^p, show that next iteration is given in the instructions

1. $$\min_{u_1} \frac{1}{2} \begin{pmatrix} u_1^T & u_2^T \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} c_1^T & c_2^T \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + d$$

<u>First Redo Matrix Expression</u>

$$\frac{1}{2} u_1^T \quad u_2^T \begin{pmatrix} H_{11} u_1 + H_{12} u_2 \\ H_{21} u_1 + H_{22} u_2 \end{pmatrix} + c_1^T u_1 + c_2^T u_2 + d$$

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

Note that H is symmetric, so $H_{21} = H_{12}^T$ thus $\frac{1}{2} u_2^T H_{21} u_1 = \frac{1}{2} u_1^T H_{12} u_2$

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

$$\frac{dV}{du_1} = H_{11} u_1 + H_{12} u_2 + c_1 = 0 \qquad u_1^* = -H_{11}^{-1} (H_{12} u_2 + c_{11}) \qquad u_p^{*1} = \begin{bmatrix} -H^{-1} (H_{12} u_2 + c_1) \\ u_2^p \end{bmatrix}$$

II) $\frac{\partial V}{\partial u_2} = u_1^T H_{12} + H_{22} u_2 + c_2 = H_{21} u_1 + H_{22} u_2 + c_2 = 0$

$$u_2^* = -H_{22}^{-1}(H_{21}u_1 + c_2)$$

$$u_p^{*2} = \begin{bmatrix} u_1^p \\ -H_{22}^{-1}(H_{21}u_1 + c_2 \end{bmatrix}$$

Convex Combination of the two:

$$\begin{bmatrix} u_1^{p+1} \\ u_2^{p+1} \end{bmatrix} = \omega_1 \begin{pmatrix} -H_{11}^{-1}(H_{12}u_2 + c_1) \\ u_2^p \end{pmatrix} + \omega_2 \begin{bmatrix} u_1^p \\ -H_{22}^{-1}(H_{21}u_1 + c_2) \end{bmatrix}$$

Now show that this takes the form $x_{p+1} = Ax + b$

$$u_1^{p+1} = -\omega_1 H_{11}^{-1} H_{12} u_2^p - \omega_1 H_{11}^{-1} c_1 + \omega_2 u_1^p$$

$$u_2^{p+1} = \omega_1 u_2^p - \omega_2 H_{22}^{-1} H_{21} u_1^p - \omega_2 H_{22}^{-1} c_2$$

Can be rewritten in the form:

$$\begin{bmatrix} u_1^{p+1} \\ u_2^{p+1} \end{bmatrix} = \begin{bmatrix} \omega_2 I & -\omega_1 H_{11}^{-1} H_{12} \\ -\omega_2 H_{22}^{-1} H_{21} & \omega_1 I \end{bmatrix} \begin{bmatrix} u_1^p \\ u_2^p \end{bmatrix} + \begin{bmatrix} -\omega_1 H_{11}^{-1} c_1 \\ -\omega_2 H_{22}^{-1} c_2 \end{bmatrix}$$

✓

b.) Establish that this is guarenteed to converge, when $|Eig(A)| < 1$

As suggested in class, use Schur Compliment

Using basic eigenvalue eqn: $\det(A - \lambda I) = 0$

$$A - \lambda I = \begin{bmatrix} (\omega_2 - \lambda) I & -\omega_1 H_{11}^{-1} H_{12} \\ -\omega_2 H_{22}^{-1} H_{21} & (\omega_1 - \lambda) I \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$\left| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right| = 0$$

Knowing that these are all matrix blocks, we can use the schur compliment to find the det()

$$\det(A - \lambda I) = \det(A_{11}) \det(A_{22} - A_{21} A_{11}^{-1} A_{12})$$

$$\det(A-\lambda I) = \det((\omega_2-\lambda)I)\det\left((\omega_1-\lambda)I + (\omega_2 H_{22}^{-1}H_{21})\left((\omega_2-\lambda)I\right)^{-1}(-\omega_1 H_{11}^{-1}H_{12})\right)$$

- $((\omega_2-\lambda)I)^{-1} = \frac{1}{\omega_2-\lambda} I$

$$= (\omega_2-\lambda)^{n_1} \det\left((\omega_1-\lambda)I + \frac{\omega_2\omega_1}{\omega_2-\lambda} H_{22}^{-1}H_{21}H_{11}^{-1}H_{12}\right) \quad H_{21} = H_{12}^T$$

$$= (\omega_2-\lambda)^{n_1} \det\left((\omega_1-\lambda)I + \frac{\omega_2\omega_1}{\omega_2-\lambda} H_{22}^{-1}H_{12}^{T}H_{11}^{-1}H_{12}\right)$$

For the $\det\left((\omega_1-\lambda)I + \frac{\omega_2\omega_1}{\omega_2-\lambda} H_{22}^{-1}H_{12}^{T}H_{11}^{-1}H_{12}\right) = 0$,

we need

$(\omega_1-\lambda)I$ & $\frac{\omega_2\omega_1}{\omega_2-\lambda} H_{22}^{-1}H_{12}^{T}H_{11}^{-1}H_{12}$ to have opposite signs.

Because $(H_{22}^{-1}H_{12}^{T}H_{11}^{-1}H_{12}) > 0$, if $\lambda > \omega_1$, that means $\lambda < \omega_2$, and if $\lambda > \omega_2$, that means $\lambda < \omega_1$, in either case, the lowest possible ω_1 or ω_2 value is zero,

So $\lambda > 0$

Redoing a different way

General Eigenvalue Eigenvector Relationship:

$$Av = \lambda v \qquad \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$w_2 v_1 - w_1 H_{11}^{-1} H_{12} v_2 = \lambda v_1$$

$$w_1 v_2 - w_2 H_{22}^{-1} H_{21} v_1 = \lambda v_2$$

$$(w_2 - \lambda) v_1 = w_1 H_{11}^{-1} H_{12} v_2$$

$$(w_1 - \lambda) v_2 = w_2 H_{22}^{-1} H_{21} v_1$$

$$\Rightarrow$$

$$v_1^T H_{11} v_1 = \frac{w_2 - \lambda}{w_1} v_1^T H_{11} H_{11}^{-1} H_{12} v_2 = \frac{w_2 - \lambda}{w_1} v_1^T H_{12} v_2$$

$$v_2^T H_{22} v_2 = \frac{w_1 - \lambda}{w_2} v_2^T H_{22} H_{22}^{-1} H_{21} v_1 = \frac{w_1 - \lambda}{w_2} v_2^T H_{21} v_1 = \frac{w_1 - \lambda}{w_2} v_1^T H_{12} v_2$$

Say we have the Quadratic Form

$$Q(v_1, v_2) = v_1^T H_{11} v_1 + v_2^T H_{22} v_2 - 2 v_1^T H_{12} v_2 = [v_1^T \; v_2^T] \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= \frac{w_2 - \lambda}{w_1} v_1^T H_{12} v_2 + \frac{w_1 - \lambda}{w_2} v_1^T H_{12} v_2 - 2 v_1^T H_{12} v_2$$

$$= \left(\frac{w_2 - \lambda}{w_1} + \frac{w_1 - \lambda}{w_2} - 2 \right) v_1^T H_{12} v_2$$

$$\frac{w_2^2 - \lambda w_2}{w_1 w_2} + \frac{w_1^2 - \lambda w_1}{w_1 w_2} - \frac{2 w_1 w_2}{w_1 w_2} = \frac{w_2^2 + w_1^2 - (w_1 + w_2)\lambda - 2 w_1 w_2}{w_1 w_2}$$

$$= \left(\frac{w_1^2 + w_2^2 - \lambda - 2 w_1 w_2}{w_1 w_2} \right) v_1^T H_{12} v_2$$

We know $Q(v_1, v_2) > 0$ (Positive Definite-ness of H)

so assuming $v_1^T H_{12} v_2 \neq 0$, we need

$$\left(\frac{w_1^2 + w_2^2 - \lambda - 2 w_1 w_2}{w_1 w_2} \right) > 0$$

$$\frac{(w_1 - w_2)^2 - \lambda}{w_1 w_2} > 0 \qquad \lambda < (w_1 - w_2)^2$$

Assuming $w_{1,2} \in (0, 1)$

$$\lambda \leq 1$$

Quadratic Fctn: $Q(v) = \begin{bmatrix} v_1^T & v_2^T \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1^T & v_2^T \end{bmatrix} \begin{bmatrix} H_{11}v_1 + H_{12}v_2 \\ H_{21}v_1 + H_{22}v_2 \end{bmatrix}$

$$= v_1^T H_{11} v_1 + v_1^T H_{12} v_2 + v_2^T H_{21} v_1 + v_2^T H_{22} v_2$$

$$= v_1^T H_{11} v_1 + 2 v_1^T H_{12} v_2 + v_2^T H_{22} v_2$$

$$\begin{bmatrix} \omega_2 I & -\omega_1 H_{11}^{-1} H_{12} \\ -\omega_2 H_{22}^{-1} H_{21} & \omega_1 I \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$\omega_2 v_1 - \omega_1 H_{11}^{-1} H_{12} v_2 = \lambda v_1 \qquad (\omega_2 - \lambda) v_1 = \omega_1 H_{11}^{-1} H_{12} v_2$

$\omega_1 v_2 - \omega_2 H_{22}^{-1} H_{21} v_1 = \lambda v_2 \qquad (\omega_1 - \lambda) v_2 = \omega_2 H_{22}^{-1} H_{21} v_1$

$$H_{12} v_2 = \frac{\omega_2 - \lambda}{\omega_1} H_{11} v_1$$

$$v_1^T H_{12} v_2 = \frac{\omega_2 - \lambda}{\omega_1} v_1^T H_{11} v_1$$

$$v_2^T H_{22} v_2 = \frac{\omega_2}{\omega_1 - \lambda} v_2^T H_{22} H_{22}^{-1} H_{21} v_1$$

$$= \frac{\omega_2}{\omega_1 - \lambda} v_2^T H_{21} v_1$$

$$= \frac{\omega_1}{\omega_1 - \lambda} v_1^T H_{12} v_2 = \frac{\omega_2}{\omega_1 - \lambda} \left(\frac{\omega_2 - \lambda}{\omega_1} \right) v_1^T H_{11} v_1$$

$Q(v) = v_1^T H_{11} v_1 + 2 v_1^T H_{12} v_2 + v_2^T H_{22} v_2$

$$= v_1^T H_{11} v_1 + 2\left(\frac{\omega_2 - \lambda}{\omega_1}\right) v_1^T H_{11} v_1 + \frac{\omega_2}{\omega_1 - \lambda}\left(\frac{\omega_2 - \lambda}{\omega_1}\right) v_1^T H_{11} v_1$$

$$= \underbrace{\left(1 + 2\left(\frac{\omega_2 - \lambda}{\omega_1}\right) + \left(\frac{\omega_2}{\omega_1 - \lambda}\right)\left(\frac{\omega_2 - \lambda}{\omega_1}\right) \right)}_{\text{This must be } > 0} v_1^T H_{11} v_1$$

$$1 + \frac{2\omega_2 - 2\lambda}{\omega_1} + \frac{\omega_2^2 - \omega_2 \lambda}{\omega_1(\omega_1 - \lambda)} = \frac{\omega_1(\omega_1 - \lambda)}{\omega_1(\omega_1 - \lambda)} + \frac{(2\omega_2 - 2\lambda)(\omega_1 - \lambda)}{\omega_1(\omega_1 - \lambda)} + \frac{\omega_2^2 - \omega_2 \lambda}{\omega_1(\omega_1 - \lambda)}$$

$$= \frac{\omega_1^2 - \omega_1 \lambda + 2\omega_1 \omega_2 - 2\omega_2 \lambda - 2\omega_1 \lambda + 2\lambda^2 + \omega_2^2 - \omega_2 \lambda}{\omega_1(\omega_1 - \lambda)}$$

$$= \omega_1^2 + \omega_2^2 + 2\omega_1\omega_2 - (\omega_1 + \omega_2)\lambda - 2(\omega_1 + \omega_2)\lambda + 2\lambda^2 > 0$$

$$= (w_1+w_2)^2 - 3\lambda + 2\lambda^2 > 0$$

$$\frac{3 \pm \sqrt{9-4(2)}}{4} = 1, \frac{1}{2} \qquad \lambda < 1$$

c.) Expressing the Converged Optimum

we know that $V(u) = \frac{1}{2} u^T H u + c^T u + d$

at the optimum, we expect $\nabla V(u) = 0$

$$\nabla V(u^*) = H u^* + c = 0$$

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

$$= \begin{bmatrix} H_{11} u_1^* + H_{12} u_2^* \\ H_{21} u_1^* + H_{22} u_2^* \end{bmatrix}$$

$$-c_1 = H_{11} u^* + H_{12} u_2^*$$

$$-c_2 = H_{21} u_1^* + H_{22} u_2^*$$

$$u_1^* = -\omega_1 H_{11}^{-1} H_{12} u_2^* - \omega_1 H_{11}^{-1} c_1 + \omega_2 u_1^*$$

$$= -\omega_1 H_{11}^{-1} H_{12} u_2^* + \omega_1 H_{11}^{-1} (H_{11} u_1^* + H_{12} u_2^*) + \omega_2 u_1^*$$

$$= -\omega_1 H_{11}^{-1} H_{12} u_2^* + \omega_1 u_1^* + \omega_1 H_{11}^{-1} H_{12} u_2^* + \omega_2 u_1^*$$

$$= (\omega_1 + \omega_2) u_1^* = u_1^*$$

$$u_2^* = \omega_1 u_2^* - \omega_2 H_{22}^{-1} H_{21} u_1^* - \omega_2 H_{22}^{-1} c_2$$

$$= \omega_1 u_2^* - \omega_2 H_{22}^{-1} H_{21} u_1^* + \omega_2 H_{22}^{-1} (H_{21} u_1^* + H_{22} u_2^*)$$

$$= \omega_1 u_2^* - \omega_2 H_{22}^{-1} H_{21} u_1^* + \omega_2 H_{22}^{-1} H_{21} u_1^* + \omega_2 u_2^*$$

$$= (\omega_1 + \omega_2) u_2^* = u_2^*$$

Q3

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

Prove that $V(u^{p+1}) < V(u^p)$

AND Size of decrease is:

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

where: $P = HD^{-1}\hat{H}D^{-1}H$ $\quad \tilde{H} = D - N \quad H = D + N$

For $D = \begin{bmatrix} \omega_1^{-1}H_{11} & 0 \\ 0 & \omega_2^{-1}H_{22} \end{bmatrix}$ & $N = \begin{bmatrix} -\omega_1^{-1}\omega_2 H_{11} & H_{12} \\ H_{21} & -\omega_1\omega_2^{-1}H_{22} \end{bmatrix}$

Step 1 (Hint), rewrite w/ origin @ $x^* = -Hc$

$$y = u - u^* = u + H^{-1}c$$

$$V(u) = V(y + u^*) = \frac{1}{2}(y+u^*)^T H (y+u^*) + c^T(y+u^*) + d$$

$$= \frac{1}{2}(y^T H y + y^T H u^* + u^{*T} H y + u^{*T} H u^*) + c^T y + c^T u^* + d$$

H is diagonal, so $y^T H u^* = u^{*T} H y$

$$= \frac{1}{2} y^T H y + u^{*T} H y + \frac{1}{2} u^{*T} H u^* + c^T y + c^T u^* + d$$

$\Downarrow$ plug in $u^* = -H^{-1}c \qquad c = -Hu^* \qquad c^T = -u^{*T}H$ ← diagonalizable

$$= \frac{1}{2} y^T H y \cancel{- c^T y} + \frac{1}{2} u^{*T} H u^* \cancel{+ c^T y} + c^T u^* + d$$

$$= \frac{1}{2} y^T H y + \left(\frac{1}{2} u^{*T} H u^* + c^T u^* + d\right) \qquad (y+u^*)^T H (y+u^*) + c(y+u^*$$

$$= \frac{1}{2} y^T H y + V(u^*)$$

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

$$= \frac{1}{2}(y^{p+1})^T H y^{p+1} - \frac{1}{2} y^{pT} H y^p$$

Step 2

turn $u^{p+1} = Au^p + B$ into y expression $\qquad y^p = u^p - u^* \qquad u^p = y^p + u^*$

$$y^{p+1} = u^{p+1} - u^* \qquad u^{p+1} = y^{p+1} + u^*$$

$$u^{p+1} = Au^p + b$$

$$y^{p+1} + u^* = A(y^p + u^*) + b$$

$$y^{p+1} = Ay^p + Au^* + b - u^*$$

Recall $A = \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix}$

$B = \begin{bmatrix} -w_1 H_{11}^{-1} c_1 \\ -w_2 H_{22}^{-1} c_2 \end{bmatrix}$

$$= \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} y^p + \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} u^* + \begin{bmatrix} -w_1 H_{11}^{-1} c_1 \\ -w_2 H_{22}^{-1} c_2 \end{bmatrix} - u^*$$

$u^* = -H^{-1}c \qquad -c = Hu^* = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} u_1^* \\ u_2^* \end{bmatrix} \qquad -c_1 = H_{11} u_1^* + H_{12} u_2^*$

$-c_2 = H_{21} u_1^* + H_{22} u_2^*$

$$\begin{bmatrix} -w_1 H_{11}^{-1} c_1 \\ -w_2 H_{22}^{-1} c_2 \end{bmatrix} = \begin{bmatrix} w_1 H_{11}^{-1} H_{11} u_1^* + w_1 H_{11}^{-1} H_{12} u_2 \\ w_2 H_{22}^{-1} H_{21} u^* + w_2 H_{22}^{-1} H_{22} u_2^* \end{bmatrix} = \begin{bmatrix} w_1 I & w_1 H_{11}^{-1} H_{12} \\ w_2 H_{22}^{-1} H_{21} & w_2 I \end{bmatrix} u^*$$

$$= \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} y^p + \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix} u^* + \begin{bmatrix} w_1 I & w_1 H_{11}^{-1} H_{12} \\ w_2 H_{22}^{-1} H_{21} & w_2 I \end{bmatrix} u^* - u^*$$

$$= \text{''} \qquad \text{''} \quad + \underbrace{\begin{bmatrix} (w_2 + w_1) I & 0 \\ 0 & (w_1 + w_2) I \end{bmatrix} u^* - u^*}_{= 0}$$

$$= Ay^p$$

$$y^{p+1} = Ay^p$$

Now Plug this into Cost Eqn

$$V(y^{p+1}) = (Ay^p)^T H (Ay^p) = \tfrac{1}{2} y^{pT} A^T H A y^p$$

Cost diff $= V(y^{p+1}) - V(y) = \tfrac{1}{2} y^{pT} A^T H A y^p - \tfrac{1}{2} y^{pT} H y^p = \tfrac{1}{2} y^{pT} (A^T H A - H) y^p$

Plug in $y^p = u^p - u^*$, $V(u) = V(y + u^*)$, $V(y) = V(y + u^*) - V(u^*)$, so $V(y) = V(u) - V(u^*)$

$$V(y^{p+1}) - V(y^p) = \left(V(v^{p+1}) - V(v^*)\right) - \left(V(v^p) - V(v^*)\right) = V(v^{p+1}) - V(v^p) = \frac{1}{2}(v^p - v^*)^T (A^T H A - H)(v^p - v^*)$$

$A^T H A - H < 0 \quad H > A^T H A\,?$ $\qquad H = N + D \qquad A^T N A + A^T D A - N - D$

$$D = \begin{bmatrix} \omega_1^{-1} H_{11} & 0 \\ 0 & \omega_2^{-1} H_{22} \end{bmatrix} \quad \& \quad N = \begin{bmatrix} -\omega_1^{-1}\omega_2 H_{11} & H_{12} \\ H_{21} & -\omega_1 \omega_2^{-1} H_{22} \end{bmatrix}$$

$$A = \begin{bmatrix} \omega_2 I & -\omega_1 H_{11}^{-1} H_{12} \\ -\omega_2 H_{22}^{-1} H_{21} & \omega_1 I \end{bmatrix}$$

$$A^T N A = \begin{bmatrix} \omega_2 I & -\omega_2 H_{22}^{-1} H_{21} \\ -\omega_1 H_{11}^{-1} H_{12} & \omega_1 I \end{bmatrix} \begin{bmatrix} -\omega_1^{-1}\omega_2 H_{11} & H_{12} \\ H_{21} & -\omega_1 \omega_2^{-1} H_{22} \end{bmatrix} \begin{bmatrix} \omega_2 I & -\omega_1 H_{11}^{-1} H_{12} \\ -\omega_2 H_{22}^{-1} H_{21} & \omega_1 I \end{bmatrix}$$

$$= \text{''} \qquad \text{''} \begin{bmatrix} -\omega_1^{-1}\omega_2^2 H_{11} - \omega_2 H_{22}^{-1} H_{12} H_{21} & \omega_2 H_{12} + \omega_1 H_{12} \\ \omega_2 H_{21} + \omega_1 H_{21} & -\omega_1 H_{11}^{-1} H_{12} H_{21} - \omega_1^2 \omega_2^{-1} H_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \omega_2 I & -\omega_2 H_{22}^{-1} H_{21} \\ -\omega_1 H_{11}^{-1} H_{12} & \omega_1 I \end{bmatrix} \begin{bmatrix} -\omega_1^{-1}\omega_2^2 H_{11} - \omega_2 H_{22}^{-1} H_{12} H_{21} & H_{12} \\ H_{21} & -\omega_1 H_{11}^{-1} H_{12} H_{21} - \omega_1^2 \omega_2^{-1} H_{22} \end{bmatrix}$$

$$= \begin{bmatrix} -\omega_1^{-1}\omega_2^3 H_{11} - \omega_2^2 H_{22}^{-1} H_{12} H_{21} - \omega_2 H_{22}^{-1} H_{21}^2 & \omega_2 H_{12} + (-\omega_2 H_{22}^{-1} H_{21})(-\omega_1 H_{11}^{-1} H_{12} H_{21} - \omega_1^2 \omega_2^{-1} H_{22}) \\ -\omega_1 H_{11}^{-1} H_{12}(-\omega_1^{-1}\omega_2^2 H_{11} - \omega_2 H_{22}^{-1} H_{12} H_{21}) + \omega_1 H_{21} & -\omega_1 H_{11}^{-1} H_{12}^2 - \omega_1^2 H_{11}^{-1} H_{12} H_{21} - \omega_1^3 \omega_2^{-1} H_{22} \end{bmatrix}$$

$$\omega_2 H_{12} + (-\omega_2 H_{22}^{-1} H_{21})(-\omega_1 H_{11}^{-1} H_{12} H_{21} - \omega_1^2 \omega_2^{-1} H_{22})$$
$$= \omega_2 H_{12} + \omega_1 \omega_2 H_{22}^{-1} H_{21} H_{11}^{-1} H_{12} H_{21} + \omega_1^2 H_{21}$$

$$-\omega_1 H_{11}^{-1} H_{12}(-\omega_1^{-1}\omega_2^2 H_{11} - \omega_2 H_{22}^{-1} H_{12} H_{21}) + \omega_1 H_{21} = \omega_2^2 H_{11}^{-1} H_{12} H_{11} + \omega_1 \omega_2 H_{11}^{-1} H_{12} H_{22}^{-1} H_{12} H_{21} + \omega_1 H_{21}$$

$$A^T N A =$$
$$\begin{bmatrix} -\omega_1^{-1}\omega_2^3 H_{11} - \omega_2^2 H_{22}^{-1} H_{12} H_{21} - \omega_2 H_{22}^{-1} H_{21}^2 & \omega_2 H_{12} + \omega_1 \omega_2 H_{22}^{-1} H_{21} H_{11}^{-1} H_{12} H_{21} + \omega_1^2 H_{21} \\ \omega_2^2 H_{11}^{-1} H_{12} H_{11} + \omega_1 \omega_2 H_{11}^{-1} H_{12} H_{22}^{-1} H_{12} H_{21} + \omega_1 H_{21} & -\omega_1 H_{11}^{-1} H_{12}^2 - \omega_1^2 H_{11}^{-1} H_{12} H_{21} - \omega_1^3 \omega_2^{-1} H_{22} \end{bmatrix}$$

$$A^\dagger O A = \begin{bmatrix} w_2 I & -w_2 H_{22}^{-1} H_{21} \\ -w_1 H_{11}^{-1} H_{12} & w_1 I \end{bmatrix} \begin{bmatrix} w_1^{-1} H_{11} & 0 \\ 0 & w_2^{-1} H_{22} \end{bmatrix} \begin{bmatrix} w_2 I & -w_1 H_{11}^{-1} H_{12} \\ -w_2 H_{22}^{-1} H_{21} & w_1 I \end{bmatrix}$$

$$= \begin{bmatrix} w_2 I & -w_2 H_{22}^{-1} H_{21} \\ -w_1 H_{11}^{-1} H_{12} & w_1 I \end{bmatrix} \begin{bmatrix} w_1^{-1} w_2 H_{11} & H_{12} \\ H_{21} & w_1 w_2^{-1} H_{22} \end{bmatrix}$$

$$= \begin{bmatrix} w_1^{-1} w_2^2 H_{11} - w_2 H_{22}^{-1} H_{21}^2 & w_2 H_{12} \\ & \end{bmatrix}$$

So $A^T H A - H = -P$ ☺

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

where $P = H D^{-1} \tilde{H} D^{-1} H$, H & D are both symmetric so can be rewritten as:

$$P = H^T D^{-1T} \tilde{H} D^{-1} H \qquad H^T D^{-1T} D D^{-1} H - H^T D^{-1T} N D^{-1} H$$

Need to show that $\tilde{H} > 0$

$$\tilde{H} = \begin{bmatrix} \omega_1^{-1} H_{11} & 0 \\ 0 & \omega_2^{-1} H_{22} \end{bmatrix} - \begin{bmatrix} -\omega_1^{-1}\omega_2 H_{11} & H_{12} \\ H_{21} & -\omega_1 \omega_2^{-1} H_{22} \end{bmatrix}$$

Both will retain positive definite-ness

$$= \begin{bmatrix} \omega_1^{-1} H_{11} + \omega_1^{-1}\omega_2 H_{11} & -H_{12} \\ -H_{21} & \omega_2^{-1} + \omega_1 \omega_2^{-1} H_{22} \end{bmatrix} = \begin{bmatrix} H_{11}\left(\frac{1+\omega_2}{\omega_1}\right) & -H_2 \\ -H_{21} & H_{22}\left(\frac{1+\omega_1}{\omega_2}\right) \end{bmatrix}$$

Positive Definite

$$= \begin{bmatrix} H_{11}\left(\frac{1+\omega_2}{1-\omega_2}\right) & -H_2 \\ -H_{21} & H_{22}\left(\frac{1+\omega_1}{1-\omega_1}\right) \end{bmatrix}$$

Because $P > 0$, $V(y^{p+1}) - V(y^p) < 0 \quad \forall\, (u^p - u^*)$

for all $u^p - u^*$ values

Q4

Change to system partitioned into 3 parts:

Considering: $V(u) = \frac{1}{2}u^T H u + c^T u$

for $H = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$, $c = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ $u_p = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

First of all, $V(u_p) = \frac{1}{2}[1 \ 0 \ 1]\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + [0 \ 1 \ 1]\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \boxed{4}$

$$u_1^{p+1} = \underset{u_1}{\operatorname{argmin}} \ \frac{1}{2}[u_1 \ 0 \ 1]\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} u_1 \\ 0 \\ 1 \end{bmatrix} + [0 \ 1 \ 1]\begin{bmatrix} u_1 \\ 0 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2}[u_1 \ 0 \ 1]\begin{bmatrix} 2u_1 + 1 \\ u_1 + 1 \\ u_1 + 2 \end{bmatrix} + 1 = \frac{1}{2}(u_1(2u_1+1) + u_1 + 2)) + 1$$

$$= \frac{1}{2}(2u_1^2 + 2u_1 + 2) + 1$$

$$= u_1^2 + u_1 + 2$$

$$\frac{dV}{du_1} = 2u_1 + 1 = 0$$

$$u_1^{p+1} = -\frac{1}{2}$$

$$u_2^{p+1} = \underset{u_2}{\operatorname{argmin}} \ \frac{1}{2}[1 \ u_2 \ 1]\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ u_2 \\ 1 \end{bmatrix} + [0 \ 1 \ 1]\begin{bmatrix} 1 \\ u_2 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2}[1 \ u_2 \ 1]\begin{bmatrix} u_2 + 3 \\ u_2 + 2 \\ u_2 + 3 \end{bmatrix} + u_2 + 1 = \frac{1}{2}\left((u_2+3) + u_2(u_2+2) + (u_2+3)\right) + u_2 + 1$$

$$= \frac{1}{2}(u_2^2 + 2u_2 + u_2 + u_2 + 3 + 3) + u_2 + 1$$

$$= \frac{1}{2}(u_2^2 + 4u_2 + 6) + u_2 + 1 = \frac{1}{2}u_2^2 + 3u_2 + 4$$

$$\frac{dV}{du_2} = u_2 + 3 = 0 \qquad u_2 = -3$$

$$u_3^{p+1} = \underset{u_3}{\operatorname{argmin}} \frac{1}{2} [1 \;\; 0 \;\; u_3] \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ u_3 \end{bmatrix} + [0 \;\; 1 \;\; 1] \begin{bmatrix} 1 \\ 0 \\ u_3 \end{bmatrix}$$

$$= \frac{1}{2} (1 \;\; 0 \;\; u_3) \begin{bmatrix} 2 + u_3 \\ 1 + u_3 \\ 1 + 2u_3 \end{bmatrix} + u_3 = \frac{1}{2}\left(2 + u_3 + u_3(1 + 2u_3)\right) + u_3$$

$$= \frac{1}{2}(2u_3^2 + 2u_3 + 2) + u_3$$

$$= u_3^2 + 2u_3 + 1$$

$$\frac{dV}{du_3} = 2u_3 + 2 = 0$$

$$u_3^{p+1} = -1$$

$$u^{p+1} = \begin{bmatrix} -\frac{1}{2} \\ -3 \\ -1 \end{bmatrix}$$

$$V(u^{p+1}) = \frac{1}{2} [-\tfrac{1}{2} \;\; -3 \;\; -1] \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ -3 \\ -1 \end{bmatrix} + [0 \;\; 1 \;\; 1] \begin{bmatrix} -\frac{1}{2} \\ -3 \\ -1 \end{bmatrix} = 6.75$$

$V(u^{p+1}) > V(u^p)$, so the three-part optimization w/o convex combination does NOT guarentee that $V(u^{p+1}) \leq V(u^p)$

✓