ME 433 – STATE SPACE CONTROL

Lecture 8

Static Optimization

1. Optimization without constraints

Problem definition: Find the values of m parameters $u_1, u_2, \ldots, u_m$ that minimize a performance function or index

$$L(u_1, u_2, \ldots, u_m) \Rightarrow dL = \frac{\partial L}{\partial u} du + \frac{1}{2} du^T \frac{\partial^2 L}{\partial u^2} du + O(3)$$

We define the decision vector $u = \begin{bmatrix} u_1 & u_2 & \ldots & u_m \end{bmatrix}^T$
and write the performance index as $L(u)$

Necessary conditions for a minimum:

$$\frac{\partial L}{\partial u} = 0 \qquad \left(\frac{\partial L}{\partial u_i} = 0, i = 1, \ldots, m \right)$$

$$\frac{\partial^2 L}{\partial u^2} \geq 0 \qquad \left(\left[\frac{\partial^2 L}{\partial u^2} \right]_{i,j} = \frac{\partial^2 L}{\partial u_i \partial u_j} \right)$$ *Positive semidefinite Hessian*

Static Optimization

Sufficient conditions for a minimum:

$$\frac{\partial L}{\partial u} = 0 \qquad \left(\frac{\partial L}{\partial u_i} = 0, i = 1, \ldots, m \right)$$

$$\frac{\partial^2 L}{\partial u^2} > 0 \qquad \left(\left[\frac{\partial^2 L}{\partial u^2} \right]_{i,j} = \frac{\partial^2 L}{\partial u_i \partial u_j} \right)$$ *Positive definite Hessian*

Note:

Positive semidefinite: $Q \geq 0 \text{ if } x^T Q x \geq 0 \quad \forall x \neq 0$

$Q \geq 0 \text{ if all } \lambda_i \geq 0, \quad Q \geq 0 \text{ if all } |m_i| \geq 0$

Positive definite: $Q > 0 \text{ if } x^T Q x > 0 \quad \forall x \neq 0$

$Q > 0 \text{ if all } \lambda_i > 0, \quad Q > 0 \text{ if all } |m_i| > 0$

λ_i: eigenvalues $\quad m_i$: principal minors

Static Optimization

Examples:

$$L = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$L = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$L = \left(u_1 - u_2^2\right)\left(u_1 - 3u_2^2\right)$$

Static Optimization

2. Optimization with constraints

Problem definition: Find the values of m parameters $u_1,\ u_2,\ldots,u_m$ that minimize a performance function or index

$$L\left(u_1,u_2,\ldots,u_m,x_1,x_2,\ldots,x_n\right)$$

Subject to the constraint equation

$$f\left(x,u\right)=0$$

The n state parameters $x_1,\ x_2,\ldots,x_n$ are determined by the decision parameters $u_1,\ u_2,\ldots,u_m$ through the constraint equation (n equations). We define:

Decision vector $\quad u=\begin{bmatrix}u_1 & u_2 & \ldots & u_m\end{bmatrix}^T$

State vector $\quad x=\begin{bmatrix}x_1 & x_2 & \ldots & x_n\end{bmatrix}^T$

Constraint vector $\quad f=\begin{bmatrix}f_1 & f_2 & \ldots & f_n\end{bmatrix}^T$

Static Optimization

If

$$L\left(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n\right)$$

and

$$f(x, u) = 0$$

are linear in both x and u, then, in general, a minimum does NOT exist. Inequalities constraints on the magnitudes of x and u are necessary to make the problem meaningful. If the inequality constraints are also linear, we are in front of a **linear programming problem**.

We will focus at the beginning on nonlinear L and f. This of course is not a guarantee of the existence of a minimum.

Static Optimization

2.1 Optimization with constraints − Approach A

At a stationary point, dL is equal to zero to first order for all increments du when df is zero, letting x change as a function of u. Thus we require

$$dL = L_x dx + L_u du = 0$$
$$df = f_x dx + f_u du = 0$$

where

$$L_x = \frac{\partial L}{\partial x}, L_u = \frac{\partial L}{\partial u}, f_x = \frac{\partial f}{\partial x}, f_u = \frac{\partial f}{\partial u}$$

Hence, if dL is zero for arbitrary du, it is *necessary* that

$$\boxed{L_u - L_x f_x^{-1} f_u = 0}$$

(m equations)

Static Optimization

2.2 Optimization with constraints − Approach B

At a stationary point, dL is equal to zero to first order for all increments du when df is zero, letting x change as a function of u. Thus we require

$$\begin{aligned} dL &= L_x dx + L_u du = 0 \\ df &= f_x dx + f_u du = 0 \end{aligned} \quad \Leftrightarrow \quad \begin{bmatrix} L_x & L_u \\ f_x & f_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} = 0$$

This set of equations defines a stationary point. For a non-trivial solution we need that the $(n+1) \times (n+m)$ matrix has rank less than $n+1$. This means that its rows must be linearly dependent. So, there exists an n vector λ (Lagrange multiplier) such that

$$\begin{bmatrix} 1 & \lambda^T \end{bmatrix} \begin{bmatrix} L_x & L_u \\ f_x & f_u \end{bmatrix} = 0 \quad \Rightarrow \quad \begin{aligned} L_x + \lambda^T f_x &= 0 \\ L_u + \lambda^T f_u &= 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} \lambda^T &= -L_x f_x^{-1} \\ \boxed{L_u - L_x f_x^{-1} f_u = 0} \end{aligned}$$

Static Optimization

2.3 Optimization with constraints − Approach C

We adjoin the constraints to the performance index to define the *Hamiltonian* function

$$H(x,u,\lambda) = L(x,u) + \lambda^T f(x,u)$$

where $\lambda \in R^n$ is a to-be-determined Lagrange multiplier. To choose x, u and λ to yield a stationary point we proceed as follows.

$$dH = \frac{\partial H}{\partial x}dx + \frac{\partial H}{\partial u}du + \frac{\partial H}{\partial \lambda}d\lambda$$

$$\frac{\partial H}{\partial \lambda} = f = 0 \qquad (n \text{ equations})$$

$$\frac{\partial H}{\partial x} = \frac{\partial L}{\partial x} + \lambda^T \frac{\partial f}{\partial x} = 0 \quad \Rightarrow \quad \lambda^T = -L_x f_x^{-1} \qquad (n \text{ equations})$$

$$\frac{\partial H}{\partial u} = \frac{\partial L}{\partial u} + \lambda^T \frac{\partial f}{\partial u} = 0 \quad \Rightarrow \quad \boxed{L_u - L_x f_x^{-1} f_u = 0} \qquad (m \text{ equations})$$

Static Optimization

2.4 Optimization with constraints − Sufficient conditions

So far, we have derived *necessary conditions* for a minimum point of $L(x,u)$ that also satisfies the constratins $f(x,u)=0$. We are interested now in *sufficient conditions*.

$$dL = \begin{bmatrix} L_x & L_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} L_{xx} & L_{xu} \\ L_{ux} & L_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + O(3) \qquad (1)$$

$$df = \begin{bmatrix} f_x & f_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xu} \\ f_{ux} & f_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + O(3) \qquad (2)$$

where

$$L_{xx} = \frac{\partial^2 L}{\partial x^2}, L_{uu} = \frac{\partial^2 L}{\partial u^2}, L_{xu} = \frac{\partial^2 L}{\partial x \partial u}, f_{xx} = \frac{\partial^2 f}{\partial x^2}, f_{uu} = \frac{\partial^2 f}{\partial u^2}, f_{xu} = \frac{\partial^2 f}{\partial x \partial u}$$

Static Optimization

$$\begin{bmatrix} 1 & \lambda^T \end{bmatrix} \begin{bmatrix} dL \\ df \end{bmatrix} = \begin{bmatrix} H_x & H_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + O(3) \quad (3)$$

For a stationary point we need $f=0$, and also that $dL=0$ to first order for all increments dx, du. Since $f=0$, we also have $df=0$. And these conditions require $H_x=0$ and $H_u=0$ (necessary conditions). By (2) we have

$$dx = -f_x^{-1} f_u du$$

Replacing this in (3) yields

$$dL = \frac{1}{2} du^T \begin{bmatrix} -f_u^T f_x^{-T} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -f_x^{-1} f_u \\ I \end{bmatrix} du + O(3)$$

Static Optimization

To ensure that this stationary point is a minimum we need dL>0 to the second order for all increments du:

$$\begin{bmatrix} -f_u^T f_x^{-T} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -f_x^{-1} f_u \\ I \end{bmatrix} > 0$$

$$H_{uu} - H_{ux} f_x^{-1} f_u - f_u^T f_x^{-T} H_{xu} + f_u^T f_x^{-T} H_{xx} f_x^{-1} f_u > 0$$

$$\left.\frac{\partial^2 L}{\partial u^2}\right|_{f=0} \equiv H_{uu} - H_{ux} f_x^{-1} f_u - f_u^T f_x^{-T} H_{xu} + f_u^T f_x^{-T} H_{xx} f_x^{-1} f_u \qquad (4)$$

Static Optimization

Examples:

(a) $$L(x,u) = \frac{1}{2}\begin{bmatrix} x & u \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ u \end{bmatrix}$$

$$f(x,u) = x - 3 = 0$$

(b) $$L(x,u) = \frac{1}{2}\left(\frac{x^2}{a^2} + \frac{u^2}{b^2}\right)$$

$$f(x,u) = x + mu - c = 0$$

(c) $$L(x,u) = \frac{1}{2}x^T Q x + \frac{1}{2}u^T R u$$

$$f(x,u) = x + Bu + c = 0$$

Static Optimization

2.5 Optimization with constraints − Lagrange multiplier

We now produce an interpretation of the Lagrange multiplier. Let us suppose that the constraints are increased by infinitesimal amounts so that we have $f(x,u)=df$, where df is an infinitesimal constant vector. How does the optimal value change?

$$dH_x^T = H_{xx}dx + H_{xu}du + f_x^T d\lambda = 0$$

$$dH_u^T = H_{ux}dx + H_{uu}du + f_u^T d\lambda = 0$$

$$df = f_x dx + f_u du$$

The partial derivatives are evaluated at the original optimal value. These equations determine dx, du, $d\lambda$.

$$dx = f_x^{-1}df - f_x^{-1}f_u du$$

$$d\lambda = -f_x^{-T}\left(H_{xx}dx + H_{xu}du\right)$$

$$du = -\left(\frac{\partial^2 L}{\partial^2 u}\right)^{-1}_{f=0}\left[H_{ux} - f_u^T f_x^{-T} H_{xx}\right]f_x^{-1}df \equiv -C\,df$$

Static Optimization

Existence of a neighboring optimal solution (for infinitesimal change in f) is guaranteed by

$$L_{uu} = \left(\frac{\partial^2 L}{\partial^2 u} \right)_{f=0} > 0$$

which is the sufficient condition for a local minimum (Equation (4)). Substituting the expression for dx and du in (3), and using $H_x = H_u = 0$, we get

$$dL = -\lambda^T df + \frac{1}{2} df^T \left[f_x^{-T} H_{xx} f_x^{-1} - C^T L_{uu} C \right] df$$

$$\frac{\partial L_{\min}}{\partial f} = -\lambda^T$$

$$\frac{\partial^2 L_{\min}}{\partial f^2} = f_x^{-T} H_{xx} f_x^{-1} - C^T L_{uu} C$$

Static Optimization

2.6 Optimization with constraints − Numerical solution

1. Select initial u

2. Determine x from $f(x,u)=0$

3. Determine λ from $\lambda^T = -L_x f_x^{-1}$

4. Determine the gradient vector $H_u = L_u + \lambda^T f_u$

5. Update the control/decision vector by $\Delta u = -kH_u$ for $k>0$ (scalar)

   (Steepest Descendent Method)

6. Determine the predicted change $\Delta L = H_u^T \Delta u = -kH_u^T H_u$. Stop if small enough. Go to step 2 otherwise.