Chapter 1: Static Optimization

1.1 Optimization without constraints

A scalar performance index function $L(u)$ is given that is a function of a control or decision vector.

$$u \in \mathbb{R}^m$$

It is desired to determine the value of u that results in a minimum value of $L(u)$

Based on the Taylor Series

$$dL = L_u^T du + \frac{1}{2} du^T L_{uu} du + O(3)$$

where O(3) represents terms of order three. The gradient of L with respect to u is the column vector

$$\boxed{L_u \triangleq \frac{\partial L}{\partial u}}$$

and the Hessian matrix is

$$\boxed{L_{uu} = \frac{\partial^2 L}{\partial u^2}}$$

L_{uu} is called the curvature matrix.

A critical or stationary point:

$$\boxed{L_u = 0}$$

For the critical point to be a local minimum, it is required that

$$\begin{aligned} dL &= L_u^T du + \frac{1}{2} du^T L_{uu} du + O(3) \\ &= \frac{1}{2} du^T L_{uu} du + O(3) \Big|_{L_u = 0} \end{aligned}$$

is positive for all increments du. This is guaranteed if the curvature matrix L_{uu} is positive definite.

$$\boxed{L_{uu} > 0}$$

If L_{uu} is negative definite, the critical point is a local maximum; and if L_{uu} is indefinite, the critical point is a saddle point.

Example 1: Quadratic Surface

Let $u \in \mathbb{R}^2$ and $Q = Q^T$

$$\begin{aligned} L(u) &= \frac{1}{2}u^T \begin{bmatrix} q_1 & q_2 \\ q_2 & q_3 \end{bmatrix} u + \begin{bmatrix} s_1 & s_2 \end{bmatrix} u \\ &\triangleq \frac{1}{2}u^T Q u + S^T u \end{aligned}$$

where $S = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$

The critical point is given by

$$L_u = \frac{\partial L}{\partial u} = Qu + S = 0$$

and the optimal control is

$$u^* = -Q^{-1}S$$

Examining the Hessian matrix

$$L_{uu} = Q$$

which determines the type of the critical point. The optimum value $L^* \triangleq L(u^*)$

$$\begin{aligned} L(u^*) &= \frac{1}{2}S^T Q^{-1} Q Q^{-1} S - S^T Q^{-1} S \\ &= \frac{1}{2}S^T Q^{-1} S - S^T Q^{-1} S \\ &= -\frac{1}{2}S^T Q^{-1} S \end{aligned}$$

Let

$$L = \frac{1}{2}u^T \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} u + \begin{bmatrix} 0 & 1 \end{bmatrix} u$$

Then

$$\begin{aligned} u^* = -Q^{-1}S &= -\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= -\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

The point u^* is minimum since $L_{uu} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ is positive definite. The minimum value is

$$L^* = -\frac{1}{2} S^T Q^{-1} S = \frac{1}{2} [0 \quad 1] \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{1}{2}$$

Example 2:

Let

$$L(u_1, u_2) = \frac{1}{2} u_1^2 + u_1 u_2 + u_2^2 + u_2$$

Where u_1 and u_2 are scalars. A critical point is present where the derivative of L with respect to all arguments are equal to zero

$$\frac{\partial L}{\partial u_1} = u_1 + u_2 = 0$$

$$\frac{\partial L}{\partial u_2} = u_1 + 2u_2 + 1 = 0$$

Solving this system of equations yields

$$u_1 = 1, \qquad u_2 = -1$$

Thus the critical point is $[1 \quad -1]^T$, the optimal value

$$L^* = L(1, -1) = \frac{1}{2} \cdot (1)^2 + (1)(-1) + (-1)^2 + (-1)$$

$$= \frac{1}{2} - 1 = -\frac{1}{2}$$

1.2 Optimization with equality constraints

Now let the scalar performance index be $L(x, u)$, a function of the control vector $u \in \mathbb{R}^m$ and auxiliary vector $x \in \mathbb{R}^n$.

The optimization problem is to determine the control vector u that minimizes $L(x, u)$ and at the same time satisfies the constraint equation:

$$f(x, u) = 0, \qquad (1.2-1)$$

Lagrange Multipliers and Hamiltonian

***Necessary conditions*:** At a stationary point, dL is equal to zero with respect to increments du when $df = 0$.

$$dL = L_u^T du + L_x^T dx = 0, \qquad (1.2-2)$$

and

$$df = f_u du + f_x dx = 0, \qquad (1.2-3)$$

First approach:
Since $(1.2-1)$ determines x for a given u, the increment dx is determined by $(1.2-3)$

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

Substituting this into $(1.2-2)$

$$dL = (L_u^T - L_x^T f_x^{-1} f_u) du = 0 \Rightarrow \boxed{L_u - f_u^T f_x^{-T} L_x = 0}$$

for any non-zero du and $f_x^{-T} = (f_x^{-1})^T$

Second approach:
Since

$$\begin{bmatrix} dL \\ df \end{bmatrix} = \begin{bmatrix} L_x^T & L_u^T \\ f_x & f_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} = 0$$

This is a homogeneous equation defines a stationary point, and it must have a non-zero solution $[dx^T \quad du^T]^T$.

The critical point is obtained if and only if

$$\begin{bmatrix} L_x^T & L_u^T \\ f_x & f_u \end{bmatrix}_{(n+1)\times(n+m)}$$

has rank less than $n+1$. That is its rows must be linearly dependent so there exist an n vector λ.

$$[1 \quad \lambda^T] \begin{bmatrix} L_x^T & L_u^T \\ f_x & f_u \end{bmatrix} = 0$$

Then

$$L_x^T + \lambda^T f_x = 0, \qquad (1.2-11)$$
$$L_u^T + \lambda^T f_u = 0, \qquad (1.2-12)$$

Solving for λ gives:

$$\lambda^T = -L_x^T f_x^{-1}, \qquad (1.2-13)$$

The vector $\lambda \in \mathbb{R}^n$ is called a Lagrange multiplier with $(1.2-13)$ and $(1.2-12)$, we have

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

<u>Third approach:</u> Define the Hamiltonian function

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

where $\lambda \in \mathbb{R}^n$ is a Lagrange multiplier.

Increments in H:

$$dH = H_x^T dx + H_u^T du + H_\lambda^T d\lambda$$

Note that

$$H_\lambda = \frac{\partial H}{\partial \lambda} = f(x,u)$$

So suppose we choose some value of u and

$$H_\lambda = 0$$

In this situation the Hamiltonian equals to

$$H|_{f=0} = L$$

Then x is determined for the given u by $f(x,u) = 0$. Avoiding solving dx in terms of du, we consider

$$H_x = 0$$

For the critical point condition, any increments in x, u and λ do not contribute to dH, so

$$\boxed{H_x = 0, \qquad H_\lambda = 0, \qquad H_u = 0}$$

$$\begin{cases} \dfrac{\partial H}{\partial \lambda} = f(x,u) = 0 \\ \dfrac{\partial H}{\partial x} = L_x + f_x^T \lambda = 0 \\ \dfrac{\partial H}{\partial u} = L_u + f_u^T \lambda = 0 \end{cases}$$

***Sufficient Conditions*:** To guarantee that the point is minimum, write Taylor series expansion for increments in L and f

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

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

To introduce the Hamiltonian, use

$$\begin{aligned} \begin{bmatrix} 1 & \lambda^T \end{bmatrix} \begin{bmatrix} dL \\ df \end{bmatrix} &= \begin{bmatrix} L_x^T & L_u^T \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + \lambda^T \begin{bmatrix} f_x & f_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} \\ &+ \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} L_{xx} & L_{xu} \\ L_{ux} & L_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} \\ &+ \lambda^T \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xu} \\ f_{ux} & f_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + O(3) \\ &= \begin{bmatrix} L_x^T + \lambda^T f_x & L_u^T + \lambda^T f_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} \\ &+ \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} L_{xx} + \lambda^T f_{xx} & L_{xu} + \lambda^T f_{xu} \\ L_{ux} + \lambda^T f_{ux} & L_{uu} + \lambda^T f_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + O(3) \\ &= \begin{bmatrix} H_x^T & H_u^T \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dx^T & du^T \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} \\ &+ O(3) \end{aligned}$$

From the necessary condition

$$H_x = 0, \qquad H_u = 0, \qquad f = 0, and\ \ df = 0$$

Then

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

Since $df = 0$

$$df = \begin{bmatrix} f_x & f_u \end{bmatrix} \begin{bmatrix} dx \\ du \end{bmatrix} + O(2) = 0 \Rightarrow dx = -f_x^{-1} f_u du + O(2)$$
$$dL = \frac{1}{2} du^T \begin{bmatrix} -f_u^T f_x^{-T} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -f_x^{-1} f_u \\ I \end{bmatrix} du + O(3)$$

To ensure a minimum, dL should be positive for all increments du. Define a curvature matrix with constant f equal to zero.

$$\boxed{\begin{aligned} L_{uu}^f \triangleq L_{uu}|_f &= \begin{bmatrix} -f_u^T f_x^{-1} & I \end{bmatrix} \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} -f_x^{-1} f_u \\ I \end{bmatrix} \\ &= H_{uu} - f_u^T f_x^{-T} H_{xu} - H_{ux} f_x^{-1} f_u + f_u^T f_x^{-T} H_{xx} f_x^{-1} f_u \end{aligned}}$$

If the curvature matrix L_{uu}^f is positive definite, then the critical point is a constrained minimum. If L_{uu}^f is negative definite, then the critical point is a constrained maximum.

Example: $\mathbf{1.2-1}$ Quadratic Surface with linear constraints
Suppose the performance index is as given:

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

Where we have simply renamed the old scalar components u_1 and u_2 as x and u respectively.

The constraints is

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

The Hamiltonian is

$$H = L + \lambda^T f = \frac{1}{2} \begin{bmatrix} x & u \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \lambda(x - 3)$$

$$= \frac{1}{2}x^2 + xu + u^2 + u + \lambda(x-3)$$

where λ is a scalar. The condition for a stationary point are

$$\begin{cases} H_\lambda = \dfrac{\partial H}{\partial \lambda} = f(x,u) = x - 3 = 0 & \Rightarrow x = 3 \\ H_x = \dfrac{\partial H}{\partial x} = x + u + \lambda = 0 \Rightarrow u + \lambda = -3 & \Rightarrow \lambda = -1 \\ H_u = \dfrac{\partial H}{\partial u} = x + 2u + 1 = 0 & \Rightarrow u = -2 \end{cases}$$

Solve the above equations, we get $x = 3$. $u = -2$ and $\lambda = -1$

The stationary point $(x, u^*) = (3, -2)$ and the value

$$L^*(3,-2) = \frac{1}{2}3^2 + 3(-2) + (-2)^2 + (-2)$$
$$= \frac{9}{2} - 6 + 4 - 2 = \frac{1}{2}$$

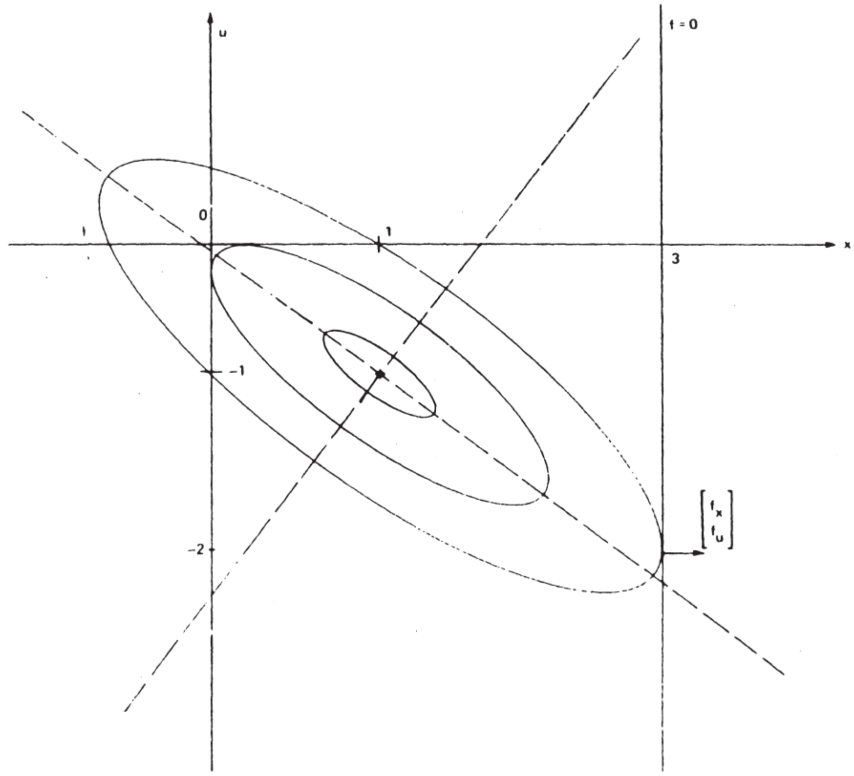
To verify that the stationary point is minimum, find the constrained curvature matrix

$$L_{uu}^f = H_{uu} - f_u^T f_x^{-T} H_{xu} - H_{ux} f_x^{-1} f_u + f_u^T f_x^{-T} H_{xx} \cdot f_x^{-1} f_u$$

where $H_{uu} = \frac{\partial^2 H}{\partial u^2} = 2,\quad f_u = \frac{\partial f}{\partial u} = 0, f_x = \frac{\partial f}{\partial x} = 1, H_{xu} = \frac{\partial H}{\partial x \partial u} = 1$, $H_{ux} = \frac{\partial H}{\partial u \partial x} = 1, H_{xx} = \frac{\partial^2 H}{\partial x^2} = 1$

$$L_{uu}^f = 2 - 0 - 1(1)^{-1} \cdot 0 + 0 = 2 > 0$$

L_{uu}^f is positive definite, the stationary point is minimum.
The contours of $L(x,u)$ and the constraints are shown in Fig. $1.2-1$.



If we look at the gradient of $f(x,u)$ in (x,u) plane.

$$\begin{bmatrix} f_x \\ f_u \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

And the gradient of $L(x,u)$ in the plane

$$\begin{bmatrix} L_x \\ L_u \end{bmatrix} = \begin{bmatrix} x+u \\ x+2u+1 \end{bmatrix}$$

At the constrained minimum $(3,-2)$

$$\begin{bmatrix} L_x \\ L_u \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Note that the gradients of f and L are parallel at the stationary point.

Example $1.2-2$ Quadratic Performance Index with Linear Constraint

Consider the quadratic performance index:

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

with linear constant

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

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $f \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, Q, R, and B are matrices and c is an n vector.
We assume $Q > 0$ and $R > 0$ (with both symmetric).

The Hamiltonian is

$$H = \frac{1}{2}x^TQx + \frac{1}{2}u^TRu + \lambda^T(x + Bu + c)$$

and the conditions for a stationary point are

$$H_\lambda = \frac{\partial H}{\partial \lambda} = x + Bu + c = 0, \qquad (1)$$

$$H_x = \frac{\partial H}{\partial x} = Qx + \lambda = 0, \qquad (2)$$

$$H_u = \frac{\partial H}{\partial u} = Ru + B^T\lambda = 0, \qquad (3)$$

Solve these equations:
From (3)

$$u = -R^{-1}B^T\lambda, \qquad (4)$$

From (2)

$$\lambda = -Qx$$

From (1)

$$\lambda = QBu + Qc$$

Using this in (4)

$$u = -R^{-1}B^T(QBu + Qc)$$

Since $R > 0$ and $B^TQB > 0$, we can invert $R + B^TQB$

$$u = -(R + B^TQB)^{-1}B^TQc$$

and the optimal state and multiplier values

$$x = -(I - B(R + B^TQB)^{-1}B^TQ)c$$

$$\lambda = (Q - QB(R + B^TQB)^{-1}B^TQ)c$$

By the matrix inversion lemma

$$\lambda = (Q^{-1} + BR^{-1}B^T)^{-1}c$$

To verify that control is a minimum

$$L_{uu}^{f} = R + B^T Q B$$

which is positive definite by restriction on R and Q

$$\begin{aligned} L^* &= \frac{1}{2} c^T [Q - QB(R + B^T QB)^{-1} B^T Q] c \\ &= \frac{1}{2} c^T \lambda \end{aligned}$$