Lecture 3a

Necessary Conditions for Finite-variable Constrained Minimization

ME260 Indian Institute of Science

Structural Optimization: Size, Shape, and Topology

G. K. Ananthasuresh

Professor, Mechanical Engineering, Indian Institute of Science, Bengaluru

suresh@iisc.ac.in

Outline of the lecture

How do constraints influence the ability to minimize the objective function?

The concept of Lagrange multipliers

Feasible space

Active and inactive constraints

Necessary conditions

What we will learn:

Constrained optimum lies on the boundary of the feasible space.

Conditions for constrained local minimum; constraint qualification

Sensitivity of the constrained optimum to small changes in constraints.

Physical meaning of Lagrange multipliers.

Two variables and an equality constraint

$$\underset{x_1, x_2}{\text{Min}} \quad f = -\, x_1 x_2$$

Subject to

$$x_1 + x_2 = 1$$

(The intent here is to maximize the product of two numbers such that their sum is equal to 1)

$$\underset{x_1}{\text{Min}} \quad f = -x_1(1 - x_1)$$

or

$$\underset{x_2}{\text{Min}} \quad f = -(1 - x_2)x_2$$

One easy way to solve this problem is to eliminate either x_1 or x_2 by expressing in terms of the other using the equality constraint and make this an unconstrained problem in one variable.

But this kind of explicit elimination of a variable may not always be possible. What do we do then?

Eliminate a variable in the first order approximation

$$\min_{x_1, x_2} f(x_1, x_2)$$

Subject to

$$h(x_1, x_2) = 0$$

Let the optimum be: (x_1^*, x_2^*)

Let the first order perturbations be:

$$\Delta x_1 = x_1 - x_1^*$$

$$\Delta x_2 = x_2 - x_2^*$$

Now, consider first-order approximations of the objective function and the constraint.

$$f(x_1, x_2) \approx f(x_1^*, x_2^*) + \left.\frac{\partial f}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 + \left.\frac{\partial f}{\partial x_2}\right|_{x_1^*, x_2^*} \Delta x_2$$

$$h(x_1, x_2) \approx h(x_1^*, x_2^*) + \left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 + \left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \Delta x_2 = \left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 + \left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \Delta x_2 = 0$$

$$\Rightarrow \Delta x_2 = -\left(\left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 \right) \Bigg/ \left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \right)$$

After eliminating one variable's first-order perturbation in terms of the other…

With $\Delta x_2 = -\left(\left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 \right) \Big/ \left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \right)$

$$f(x_1, x_2) \approx f(x_1^*, x_2^*) + \left.\frac{\partial f}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 - \left.\frac{\partial f}{\partial x_2}\right|_{x_1^*, x_2^*} \frac{\left(\left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 \right)}{\left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \right)} = f(x_1^*, x_2^*) + \left.\frac{\partial f}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1 + \lambda \left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \Delta x_1$$

where $\lambda = -\frac{\left(\left.\frac{\partial f}{\partial x_2}\right|_{x_1^*, x_2^*} \right)}{\left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \right)}$

$$\left(\left.\frac{\partial f}{\partial x_1}\right|_{x_1^*, x_2^*} \right) + / \left(\left.\frac{\partial h}{\partial x_1}\right|_{x_1^*, x_2^*} \right) = 0$$

$$\left(\left.\frac{\partial f}{\partial x_2}\right|_{x_1^*, x_2^*} \right) + / \left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*, x_2^*} \right) = 0$$

Because the first order derivative should be zero for any dx_1

Some generality and a new concept

Two things are remarkable about the two equations we got: (i) they are similar and (ii) they both give a similar expression for the constant, **/** .

$$\left(\left.\frac{\partial f}{\partial x_1}\right|_{x_1^*,x_2^*}\right) + / \left(\left.\frac{\partial h}{\partial x_1}\right|_{x_1^*,x_2^*}\right) = 0$$

$$\left(\left.\frac{\partial f}{\partial x_2}\right|_{x_1^*,x_2^*}\right) + / \left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*,x_2^*}\right) = 0$$

$$/ = - \frac{\left(\left.\frac{\partial f}{\partial x_1}\right|_{x_1^*,x_2^*}\right)}{\left(\left.\frac{\partial h}{\partial x_1}\right|_{x_1^*,x_2^*}\right)} = - \frac{\left(\left.\frac{\partial f}{\partial x_2}\right|_{x_1^*,x_2^*}\right)}{\left(\left.\frac{\partial h}{\partial x_2}\right|_{x_1^*,x_2^*}\right)}$$

/ This is called the Lagrange multiplier corresponding to the equality constraint.

Think about what the Lagrange multiplier **physically** means…
It is the negative of the ratio of the rate of change of objective function to the rate of change of the constraint with respect to either variable.

The concept of the Lagrangian

$$\underset{x_1,x_2}{\text{Min}} \; f(x_1,x_2)$$

Subject to

$$h(x_1,x_2)=0$$

An alternative formulation…

$$\underset{x_1,x_2}{\text{Min}} \; L = f(x_1,x_2) + / \; h(x_1,x_2)$$

Both have the same necessary conditions!

$$\left(\left. \frac{\partial f}{\partial x_1} \right|_{x_1^*,x_2^*} \right) + / \left(\left. \frac{\partial h}{\partial x_1} \right|_{x_1^*,x_2^*} \right) = 0$$

$$\left(\left. \frac{\partial f}{\partial x_2} \right|_{x_1^*,x_2^*} \right) + / \left(\left. \frac{\partial h}{\partial x_2} \right|_{x_1^*,x_2^*} \right) = 0$$

L is called the Lagrangian.

General problem in two variables and one equality constraint

$$\min_{x_1,x_2} f(x_1,x_2)$$

*S*ubject to

$$h(x_1,x_2)=0$$

$$\Rightarrow \quad \min_{x_1,x_2} L = f(x_1,x_2)+\lambda\, h(x_1,x_2)$$

Necessary conditions

$$\frac{\partial f}{\partial x_1}+\lambda\frac{\partial h}{\partial x_1}=0$$

$$\frac{\partial f}{\partial x_2}+\lambda\frac{\partial h}{\partial x_2}=0$$

$$h(x_1,x_2)=0$$

Three variables (x_1,x_2,λ)
And three equations!

We are fine.

n variables and m equality constraints

$$\min_{x_1, x_2, \cdots, x_n} \quad f(x_1, x_2, \cdots, x_n)$$

Subject to

$$h_1(x_1, x_2, \cdots, x_n) = 0$$
$$h_2(x_1, x_2, \cdots, x_n) = 0$$
$$\cdots$$
$$h_m(x_1, x_2, \cdots, x_n) = 0$$

Short form

$$\min_{\mathbf{x}} \quad f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \text{ and } \mathbf{h} = \begin{Bmatrix} h_1(x_1, x_2, \cdots, x_n) \\ h_2(x_1, x_2, \cdots, x_n) \\ \vdots \\ h_m(x_1, x_2, \cdots, x_n) \end{Bmatrix}$$

Can there be more constrains than variables?

$m > n$ No; feasible values may not exist. It is over-constrained.

$m = n$ Some discrete feasible values may exist; but cannot do minimization.

$m < n$ This must be true in order to do minimization of the objective function.

Feasible space; partitioning variables

If there are m constraints (note: $m < n$), we can choose only (n-m) variables freely because m variables can be found using the m equality constraints.

So, we search in the (n-m)-dimensional feasible space.

Feasible space is the reduced space that satisfies the constraints.

When we say it is (n-m)-dimensional, we mean that m variables are somehow eliminated using m equality constraints.

Let us partition n variables in to **s** (solution or dependent) variables and **d** (decision or independent variables).

$$\mathbf{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{Bmatrix} = \begin{Bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \\ d_1 \\ d_2 \\ \vdots \\ d_{n-m} \end{Bmatrix} = \begin{Bmatrix} \mathbf{s} \\ \mathbf{d} \end{Bmatrix}$$

Taylor series again

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \sum_{i=1}^{n} \frac{\partial f(\mathbf{x}^*)}{\partial x_i}(x_i - x_i^*) + O(2)$$

$$\approx f(\mathbf{x}^*) + \nabla f^T(\mathbf{x}^*)\Delta\mathbf{x}^*$$ Approximated to first order.

$$\approx f(\mathbf{x}^*) + \nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\Delta\mathbf{s}^* + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*)\Delta\mathbf{d}^*$$ As per partitioned variables.

For the necessary condition, we want the first order terms go to zero; then, the function value does not change in the vicinity of the minimum up to first order.
But we know that perturbation in **s** variables cannot be independent of those in **d** variables. So… (next slide)

Taylor series for equality constraints

$$h_j(\mathbf{x}) = h_j(\mathbf{x}^*) + \sum_{i=1}^{n} \frac{\partial h_j(\mathbf{x}^*)}{\partial x_i}(x_i - x_i^*) + O(2) \qquad j = 1, 2, \cdots, m$$

$$\approx h_j(\mathbf{x}^*) + \nabla h_j^T(\mathbf{x}^*) \mathrm{D}\mathbf{x}^*$$

As per partitioned variables.

$$\approx h_j(\mathbf{x}^*) + \nabla_{\mathbf{s}} h_j^T(\mathbf{x}^*) \mathrm{D}\mathbf{s}^* + \nabla_{\mathbf{d}} h_j^T(\mathbf{x}^*) \mathrm{D}\mathbf{d}^* = 0$$

($h_j(\mathbf{x}^*)$ is struck through, marked 0.)

Because $\mathbf{x}^*$ is feasible; i.e., it satisfies the constraints.

$$\Rightarrow \nabla_{\mathbf{s}} h_j^T(\mathbf{x}^*) \mathrm{D}\mathbf{s}^* + \nabla_{\mathbf{d}} h_j^T(\mathbf{x}^*) \mathrm{D}\mathbf{d}^* = 0$$

Because $\mathbf{x}$ should remain feasible after perturbation from $\mathbf{x}^*$.

Eliminating s-perturbations

$$\nabla_{\mathbf{s}} h_j^T(\mathbf{x}^*)\mathrm{D}\mathbf{s}^* + \nabla_{\mathbf{d}} h_j^T(\mathbf{x}^*)\mathrm{D}\mathbf{d}^* = 0 \qquad j = 1, 2, \cdots, m$$

$$\underbrace{\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*)}_{m \times m} \underbrace{\Delta \mathbf{s}^*}_{m \times 1} + \underbrace{\nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*)}_{m \times (n-m)} \underbrace{\Delta \mathbf{d}^*}_{(n-m) \times 1} = \underbrace{\mathbf{0}}_{m \times 1}$$

Compact notation for all constraints.

Note the sizes of the quantities.

$$\Rightarrow \mathrm{D}\mathbf{s}^* = -\left[\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*)\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*)\mathrm{D}\mathbf{d}^*$$

Remember that we can take the inverse of the m x m matrix here only if it is not singular. This implies that the gradients of the constraints should be linearly independent. This is known as constraint qualification.

Reduced gradient

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\mathbf{Ds}^* + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*)\mathbf{Dd}^*$$

Should be zero for the necessary condition for a minimum.

With $\mathbf{Ds}^* = -\left[\nabla_{\mathbf{s}}\mathbf{h}^T(\mathbf{x}^*)\right]^{-1}\nabla_{\mathbf{d}}\mathbf{h}^T(\mathbf{x}^*)\mathbf{Dd}^*$

We get $\nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\mathbf{Ds}^* + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*)\mathbf{Dd}^* = 0$

$$-\nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\left[\nabla_{\mathbf{s}}\mathbf{h}^T(\mathbf{x}^*)\right]^{-1}\nabla_{\mathbf{d}}\mathbf{h}^T(\mathbf{x}^*)\mathbf{Dd}^* + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*)\mathbf{Dd}^* = 0$$

$$\underbrace{\left\{-\nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\left[\nabla_{\mathbf{s}}\mathbf{h}^T(\mathbf{x}^*)\right]^{-1}\nabla_{\mathbf{d}}\mathbf{h}^T(\mathbf{x}^*) + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*)\right\}}_{\nabla z(\mathbf{d})}\mathbf{Dd}^* = 0$$

Reduced gradient in the p=n-m space.

Think of *f* as *z* that depends only the **d** variables

The multipliers, again

Reduced gradient is zero is the necessary condition.

$$\left\{ -\nabla_s f^T(\mathbf{x}^*) \left[\nabla_s \mathbf{h}^T(\mathbf{x}^*) \right]^{-1} \nabla_d \mathbf{h}^T(\mathbf{x}^*) + \nabla_d f^T(\mathbf{x}^*) \right\} = \mathbf{0}$$

$$\boldsymbol{\lambda} \nabla_d \mathbf{h}^T(\mathbf{x}^*) + \nabla_d f^T(\mathbf{x}^*) = \mathbf{0}$$

Where we used the new symbol for

$$\boldsymbol{\lambda} = -\nabla_s f^T(\mathbf{x}^*) \left[\nabla_s \mathbf{h}^T(\mathbf{x}^*) \right]^{-1}$$

$$\boldsymbol{\lambda} \nabla_d \mathbf{h}^T(\mathbf{x}^*) + \nabla_d f^T(\mathbf{x}^*) = \mathbf{0}$$

$$\boldsymbol{\lambda} \nabla_s \mathbf{h}^T(\mathbf{x}^*) + \nabla_s f^T(\mathbf{x}^*) = \mathbf{0}$$

Lagrange multipliers appear again.
Compare with the two-variable case on slide 5 of this lecture.
Same story here!

Notice that both equations have the same form; one is gradient w.r.t. to **d** and the other w.r.t. **s**.

The Lagrangian

$$\boldsymbol{\lambda}\nabla_{\mathbf{d}}\mathbf{h}^T(\mathbf{x}^*)+\nabla_{\mathbf{d}}f^T(\mathbf{x}^*)=\mathbf{0}$$

$$\boldsymbol{\lambda}\nabla_{\mathbf{s}}\mathbf{h}^T(\mathbf{x}^*)+\nabla_{\mathbf{s}}f^T(\mathbf{x}^*)=\mathbf{0}$$

$$\boldsymbol{\lambda}\nabla_{\mathbf{x}}\mathbf{h}^T(\mathbf{x}^*)+\nabla_{\mathbf{x}}f^T(\mathbf{x}^*)=\mathbf{0}$$

With $L = f + \boldsymbol{\lambda}\mathbf{h}$

$$L = f + \sum_{j=1}^{m}\lambda_j h_j$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}^*)\boldsymbol{\lambda}^T = \mathbf{0}$$

$\boldsymbol{\lambda}_{1\times m}$

is a row vector.

$$\Rightarrow \frac{\partial f}{\partial x_i} + \sum_{j=1}\lambda_j \frac{\partial h_j}{\partial x_i} = 0 \qquad i = 1, 2, \cdots, n$$

We have n scalar equations here.

Necessary conditions for equality-constrained minimization problem

$$\underset{\mathbf{x}}{\text{Min}} \;\; f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*) \boldsymbol{\lambda}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}$$

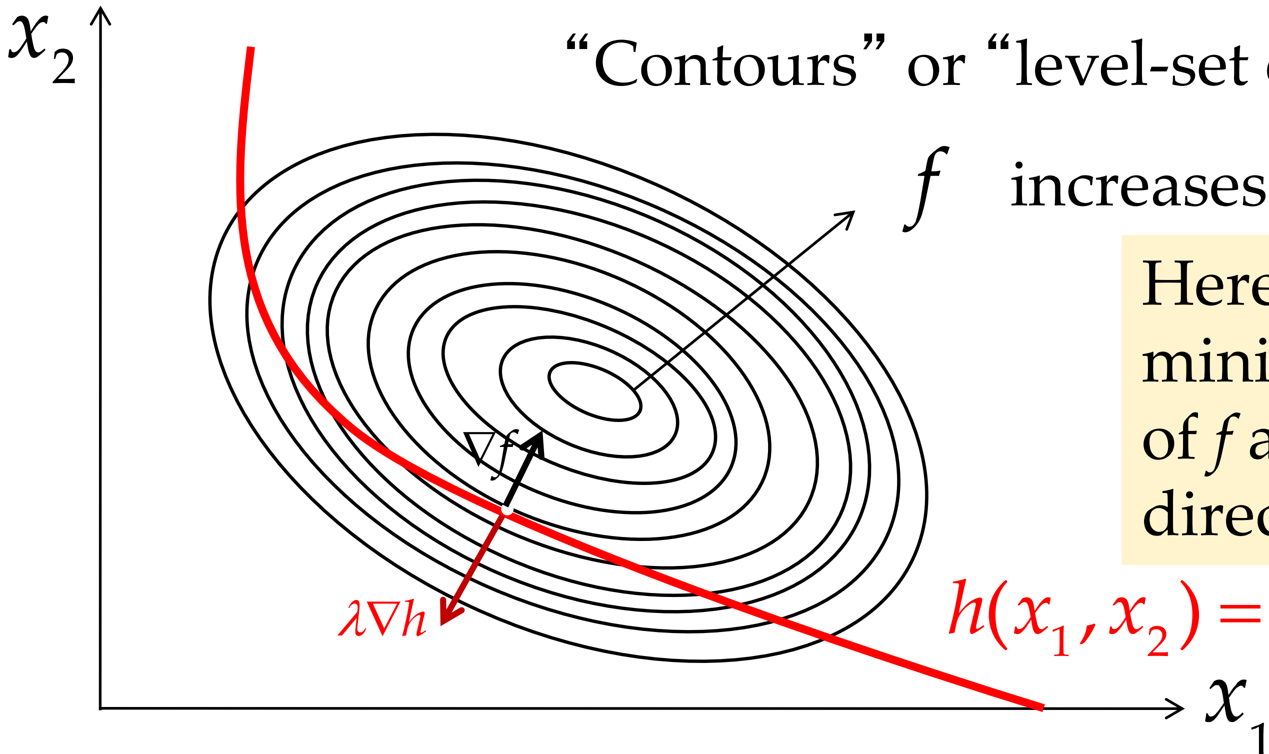
Variables: $n+m$
Equations: $n+m$
So, we are fine.

Geometric interpretation

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*) \boldsymbol{\lambda}^T = \mathbf{0}$$

$$\Rightarrow \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial h_j}{\partial x_i} = 0 \qquad i = 1, 2, \cdots, n$$

This means that the gradient of the objective function is a linear combination of the gradients of the constraints.



Here, at the constrained minimum, the gradients of *f* an *h* are in the same direction.

With inequality constraints

$$\underset{\mathbf{x}}{\text{Min}}\ \ f(\mathbf{x})$$

*S*ubject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \pounds \mathbf{0}$$

p inequality constraints.

$$g_k(\mathbf{x}) \pounds 0 \quad k = 1, 2, \cdots, p$$

Inequality constraints may be **active** or **inactive** at the minimum point.

$$g = 0 \qquad g < 0$$

Active constraints should be treated just like equality constraints.

Inequality constraints may simply be ignored because they don't not play any role.

Complementarity conditions

How do we know if an inequality constraint is active or not?
We don't.
So, we express it in the form of equations!

$$m_k g_k = 0 \quad k = 1, 2, \cdots, p$$

Lagrange multiplier corresponding to k[th] inequality constraint

This is an interesting set of equations:

Either multiplier is zero or the constraint is zero.

They are called complementarity conditions.

Both may also be zero under special situations.

Necessary conditions for constrained minimization with equalities and inequalities.

$$\underset{\mathbf{x}}{\text{Min}} \quad f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \text{ £ } \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*) \boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*) \boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \leq 0$$

$$\mu_k \, g_k(\mathbf{x}^*) = 0; \quad k = 1, 2, \cdots, p$$

Variables: $n+m+p$

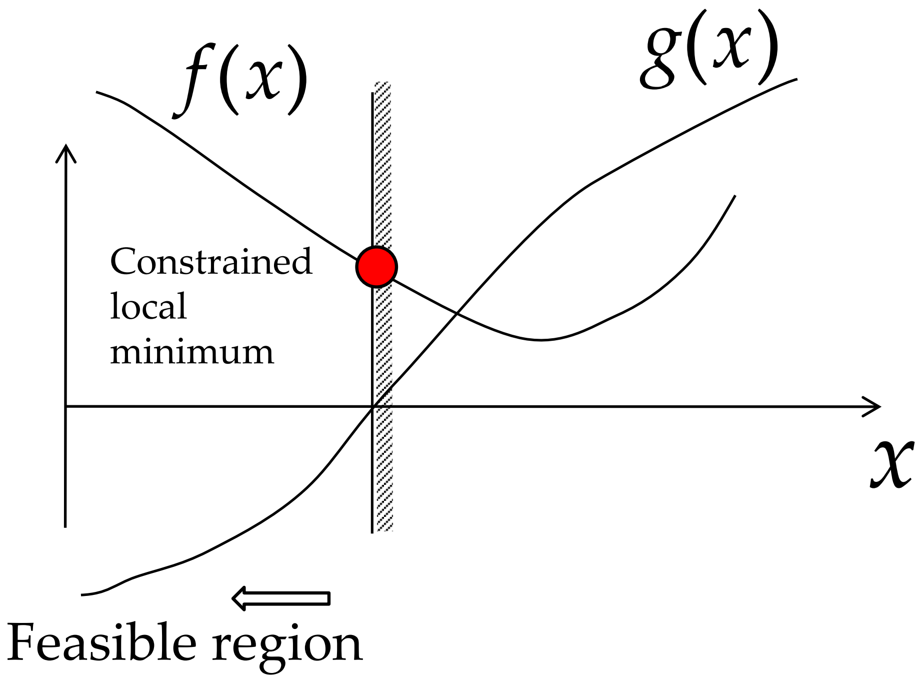
Equations: $n+m+p$

So, we are fine.

But we are not done yet.

The Lagrange multipliers of inequality constraints are restricted in sign. Let us discuss why.

Inequality constraint; simplest problem



$$f(x\ ) = f(x^*) + \frac{df(x^*)}{dx}\Delta x^* + O(2)$$

$$f'(x^*)\Delta x^* \ \geq \ 0$$

For x* to be a local minimum.

$$g(x\ ) = g(x^*) + \frac{dg(x^*)}{dx}\Delta x^* + O(2)$$

$$\Rightarrow g'(x^*)\Delta x^* \ \leq \ 0$$

For x* to continue to satisfy the inequality even after perturbation.

Minimizability vs. feasibility

$f'(x^*)\Delta x^* \geq 0$ $\qquad$ $g'(x^*)\Delta x^* \leq 0$

But necessary condition requires: $f'(x^*) + \mu g'(x^*) = 0$

Multiply both sides by Δx^* : $\underbrace{f'(x^*)\Delta x^*}_{\geq 0} + \mu \underbrace{g'(x^*)\Delta x^*}_{\leq 0} = 0$

It is a simple but a good explanation for the non-negativity of the Lagrange multipliers of inequality constraints.

So, this has to be non-negative. That is, $\mu \geq 0$

Necessary condition for general constrained minimization problem

$$\underset{\mathbf{x}}{\text{Min}} \quad f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \text{ £ } \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*)\boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*)\boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \leq 0$$

$$\mu_k\, g_k(\mathbf{x}^*) = 0; \;\; \mu_k \geq 0; \;\; k = 1, 2, \cdots, p$$

Variables: $n+m+p$

Equations: $n+m+p$

Number of inequalities: p

The first condition follows from the Lagrangian.

$$L = f(\mathbf{x}) + \boldsymbol{\lambda}\,\mathbf{h}(\mathbf{x}^*) + \boldsymbol{\mu}\,\mathbf{g}(\mathbf{x}^*)$$

Necessary conditions: KKT conditions

$$\underset{\mathbf{x}}{\text{Min}} \;\; f(\mathbf{x})$$

*S*ubject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$\mathbf{g}(\mathbf{x})$ £ $\mathbf{0}$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*)\boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*)\boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \leq 0$$

$$\mu_k \, g_k(\mathbf{x}^*) = 0; \;\; \mu_k \geq 0; \;\; k = 1, 2, \cdots, p$$

Karush-Kuhn-Tucker conditions.

Independently done at Princeton University; later Kuhn dug up Karush's work and gave credit that is due to him. A rare and admirable gesture.

Had done this in his master's thesis at University of Chicago before Kuhn and Tucker.

A caveat: constraint qualification

KKT conditions are applicable as "necessary conditions" only if the constraints qualification is satisfied.

Constraint qualification requires that the gradients of the equality constraints and active inequality constraints be linearly independent at the optimum.

- See slide 13.
- One can construct special example where a point is a minimum but KKT conditions are not satisfied.
  - How can "necessary" conditions be not satisfied?
  - It is because at such special points "constraint qualification" is not satisfied. So, KKT conditions are not applicable.

The end note

Necessary conditions
for finite-variable constrained optimization

- Two variables and one equality constraint
  The concept of Lagrange multiplier and the Lagrangian
- Feasible space
  Reduced gradient with equality constraints
  Lagrange multipliers
- Constraint qualification
- Inequality constraint and the implication of the sign of the Lagrange multiplier
  Complementarity conditions
- **Karush-Kuhn-Tucker necessary conditions**

Thanks