# **Urban Transportation Networks**
# Yossi Sheffi, Prentice Hall,1985
# Chapter2 “Basic Concepts in Minimization Problems”


2015/4/24 （金）
理論勉強会2015#2
B4　庄司惟

# 非線形最適化

・こういうやつ

$$\min f(x_1, x_2, \dots, x_n)$$

目的関数
Objective function

*subject to*

$$g_1(x_1, x_2, \dots, x_n) \geq 0$$
$$g_2(x_1, x_2, \dots, x_n) \geq 0$$
$$\vdots$$
$$g_J(x_1, x_2, \dots, x_n) \geq 0$$

制約条件
Constraint

・均衡配分問題では、各経路・リンクの交通量を変数として、旅行時間の和などを最適化する。

→ Part 1
Part 2
Part 3
Part 4
Part 5

# 本日の目標

・制約なしの非線形最適化を考える

・その後、制約条件付の実用的な非線形最適化を考える

・また、イメージを掴み易いよう、１変数から導入して多変数問題へと拡張します。

# いままでは…

・制約無し
→とりあえず微分。
→確かめるときは二階微分。

・制約条件付
Lagrangeの未定乗数法（等式制約）
KKT(Karush-Kuhn-Tucker)条件（不等式制約）
→とりあえずそのとおり計算。

※ではなく、本質的な理解をすることが目的。

# 制約条件なし①

・制約条件無し、一次元最適化問題

※以降、最適解を$x^*$と表す。

<u>一次の最適性条件→候補を絞る</u>

$$\frac{df(x^*)}{dx} = 0$$

（上式を満たす点を、停留点＝stationary point と呼びます。）

<u>二次の最適性条件→極小値なのかを判断</u>

$$\frac{d^2f(x^*)}{dx^2} > 0$$

Part 1

→ **Part 2**

Part 3

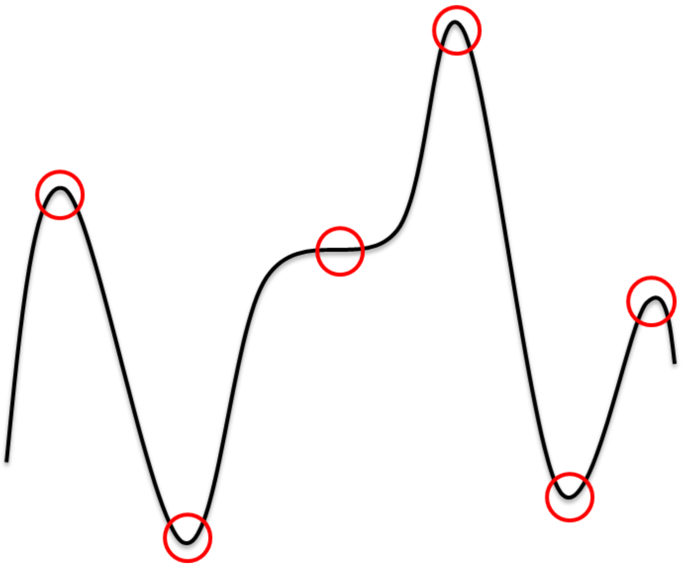
Part 4

Part 5

# 制約条件なし②

<u>一次の最適性条件→候補を絞る</u>

$$\frac{df(\bar{x})}{dx} = 0$$



・極小のほかにも、
極大or変曲の可能性ある。

→停留点が局所的最適解であるための条件は？
→二次の最適性条件

Part 1

→ **Part 2**

Part 3

Part 4

Part 5

# 制約条件なし③

<u>二次の最適性条件→極小値なのかを判断</u>

「停留点近傍で関数が狭義凸関数であること」

# 制約条件なし④

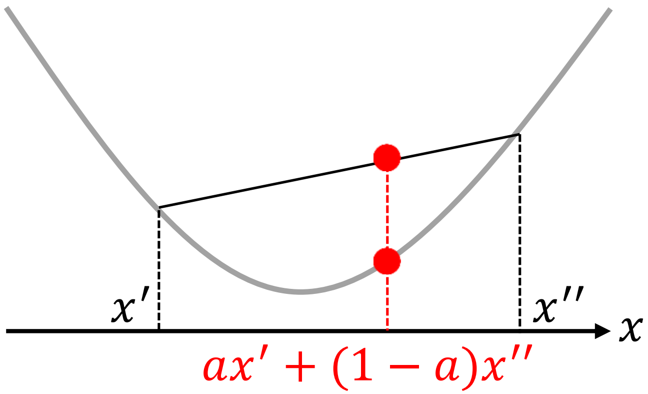
狭義凸関数（strictly convex function）
…関数上の任意の二点を結ぶ線分がつねに関数の上部にあること。
…下のような三種類の同値条件で表せます。
①$f[ax' + (1-a)x''] < af(x') + (1-a)f(x'')$
②$f(x') + \frac{df(x')}{dx}(x''' - x') < f(x''')$
③$\frac{d^2f(x)}{dx^2}$>0

（つまり、①～③のどれか一つが満たされていればよいことになります。）
（各条件に等号を許せば、そのとき関数は単に凸関数であると言います。）

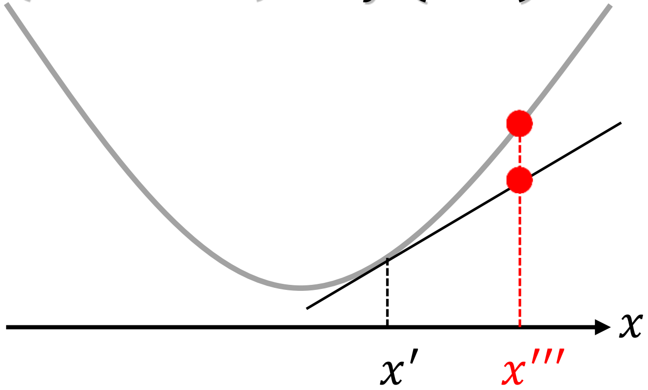
# 制約条件なし⑤

**狭義凸関数**（strictly convex function）

①$f[ax' + (1-a)x''] < af(x') + (1-a)f(x'')$



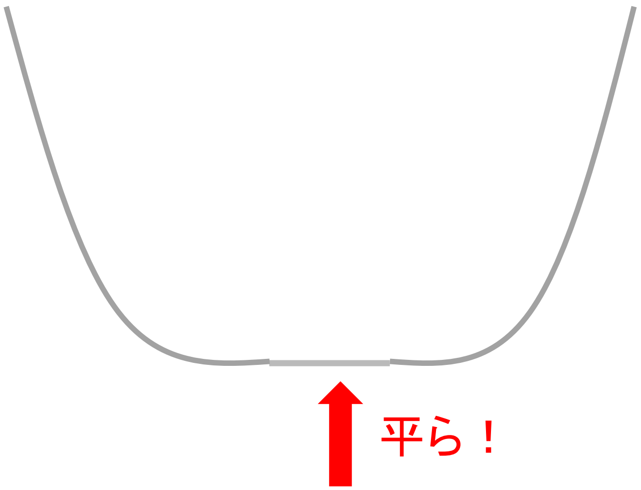

②$f(x') + \frac{df(x')}{dx}(x''' - x') < f(x''')$



Part 1
→ Part 2
Part 3
Part 4
Part 5

# 制約条件なし⑥

単なる凸関数は条件①～③の等号も許すので、
下みたいなのも凸関数です。



# 制約条件なし⑦

・一般に、極小値を二つ以上持つ関数から唯一の最小値をみつけるのは容易ではない。

・１変数関数では「極小値を抽出して比較すれば簡単」と思うかもしれないが、多変数関数ではそのような手法は困難である。

・利用者均衡モデルは多変数関数になる。

・ゆえに、計算可能な、「目的関数が一つの極小値しか持たない関数」の性質を知り、そのような関数を利用するようにしなければならない。

# 制約条件なし⑧

一次の最適性条件
→停留点を求める行為

二次の最適性条件
→狭義凸関数であるかどうかを調べる行為

といえるでしょう。

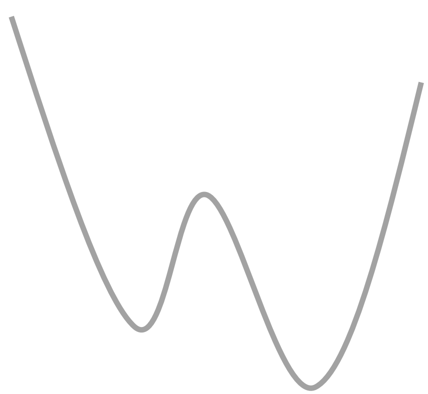
この双方を満たす解を、
局所的最適解(local Optimal Solution)
と呼びます。

# 制約条件なし⑨

一次と二次の最適性条件を満たす**局所的最適解**が**大域的最適解(Global Optimal Solution)**であることを示すには？

「関数全体が**凸関数**であれば良い。」

(×)



(○)

# 制約条件なし⑩

以上をまとめると、１変数制約無しの場合、

・一次の最適性条件
→**停留点**
・二次の最適性条件
→**停留点＝極小点**であることを保証
・関数が凸関数
→**極小点＝大域的最適解**であることを保証

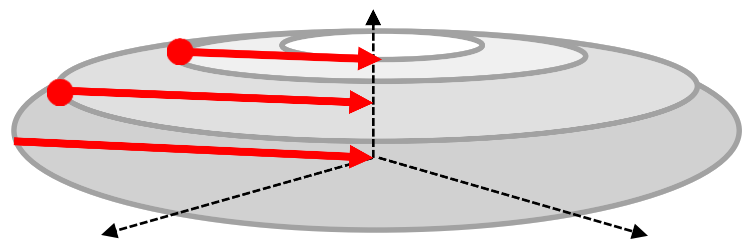
Part 1
→ **Part 2**
Part 3
Part 4
Part 5

# 制約条件なし⑪
# 多次元問題への拡張

一階微分→gradientになる。

$$\nabla f(\vec{x}) = \left(\frac{\partial f(\vec{x})}{\partial x_1}, \dots, \frac{\partial f(\vec{x})}{\partial x_n}\right)$$

※目的関数が一番大きく増加する向きに伸びるベクトル。



<u>一次の最適性条件</u>

→停留点では$\nabla f\left(\overrightarrow{x^*}\right) = \vec{0}$である。

<u>二次の最適性条件</u>

→？？

# 制約条件なし⑫
# 多次元問題への拡張

二次の最適性条件

＝局所的最適解であることの保証

＝目的関数が解近傍で狭義凸関数であることの保証

三つの同値条件のうち①②はそのまま理解可能。

①$f\left[a\overrightarrow{x'}+(1-a)\overrightarrow{x''}\right]<af\left(\overrightarrow{x'}\right)+(1-a)f\left(\overrightarrow{x''}\right)$

②$f\left(\overrightarrow{x'}\right)+\nabla f\left(\overrightarrow{x'}\right)\left(\overrightarrow{x'''}-\overrightarrow{x'}\right)<f\left(\overrightarrow{x'''}\right)$

では、二階微分の条件③$\frac{d^2f(x)}{dx^2}$>0はどう表されるのか。

Part 1
→ Part 2
Part 3
Part 4
Part 5

# 制約条件なし⑬
# 多次元問題への拡張

（結論）

「Hesse行列$\{\nabla^2 f(\overrightarrow{x^*})\}$が正定値である」とき目的関数は狭義凸関数。

$$\nabla^2 f(\overrightarrow{x^*}) = \begin{bmatrix} \frac{\partial^2 f(\overrightarrow{x^*})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\overrightarrow{x^*})}{\partial x_1 \partial x_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\overrightarrow{x^*})}{\partial x_M \partial x_1} & \cdots & \frac{\partial^2 f(\overrightarrow{x^*})}{\partial x_M^2} \end{bmatrix}$$

が、$\vec{0}$でない任意のベクトル$\vec{Z}$に対して、

$\vec{Z}^T\{\nabla^2 f(\overrightarrow{x^*})\}\vec{Z} > \lambda$となること。（$\nabla^2 f(\overrightarrow{x^*})$が正定値）

→どう理解したらよいか。

# 制約条件なし⑭
# 多次元問題への拡張

・Hesse行列が正定値でないと、点$\overrightarrow{x^*}$からある点$\overrightarrow{x^*}+a\vec{d}$（$a$は十分小さな値、$\vec{d}$は移動する方向）に移動すると目的関数の値が小さくなることがある。

<軽い証明>

$\nabla^2 f$の連続性より、点$\overrightarrow{x^*}$のある近傍$N\left(\overrightarrow{x^*},\hat{r}\right)=\{\vec{x}|\left\|\overrightarrow{x^*}-\overrightarrow{x^*}\right\|<,\hat{r}\}$が存在して、

$\nabla^2 f(\vec{x})$も正定値（$\forall x\in N\left(\overrightarrow{x^*},\hat{r}\right)$）になる。

一方、その近傍におさまる範囲で$\overrightarrow{x^*}+\vec{d}$の$\vec{d}$（$\neq\vec{0}$）を与えると、

Taylorの定理より

$$f\left(\overrightarrow{x^*}+\vec{d}\right)=f\left(\overrightarrow{x^*}\right)+\left\langle\nabla f\left(\overrightarrow{x^*}\right),\vec{d}\right\rangle+\frac{1}{2}\left\langle\nabla^2 f\left(\overrightarrow{x^*}+t\vec{d}\right)\vec{d},\vec{d}\right\rangle$$
$$=f\left(\overrightarrow{x^*}\right)+\frac{1}{2}\left\langle\nabla^2 f\left(\overrightarrow{x^*}+t\vec{d}\right)\vec{d},\vec{d}\right\rangle$$

を満たす$t(\in(0,1))$が存在する。

$\overrightarrow{x^*}+t\vec{d}\in N\left(\overrightarrow{x^*},\hat{r}\right)$より、$\nabla^2 f\left(\overrightarrow{x^*}+t\vec{d}\right)$もまた正定値であるから、

$\frac{1}{2}\left\langle\nabla^2 f\left(\overrightarrow{x^*}+t\vec{d}\right)\vec{d},\vec{d}\right\rangle>0$であり、$f\left(\overrightarrow{x^*}+\vec{d}\right)>f\left(\overrightarrow{x^*}\right)$がわかる。■

# 制約条件有り①

- 制約条件付の１変数問題
- 制約条件を付加した最適化問題の一般的表現

目的関数：$f(x)$ →最小

制約条件：$g_i(x) \geq 0 (i = 1, \dots, m)$

- 何が変わる？

Part 1
Part 2
→ **Part 3**
Part 4
Part 5

# 制約条件有り②

・制約条件有りのもとでは、
①実行可能領域の内側に最適解が存在する。
②実行可能領域の縁（制約条件式の境界上）に最適解が存在する。
の二種類のパターンを想定すればよい。

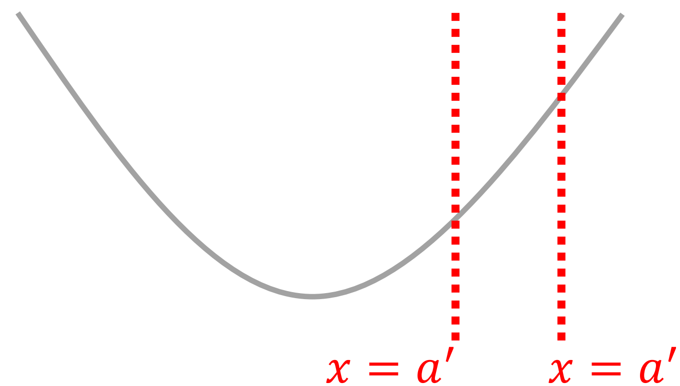
①→一次の最適性条件（一階微分＝０）がそのまま適用できる。
②→そうはいかない…

# 制約条件有り③ sheffiの記述 “same sign” condition

制約条件$g_1(\bar{x}) = x - a' > 0, g_2(\bar{x}) = -x + a'' > 0$

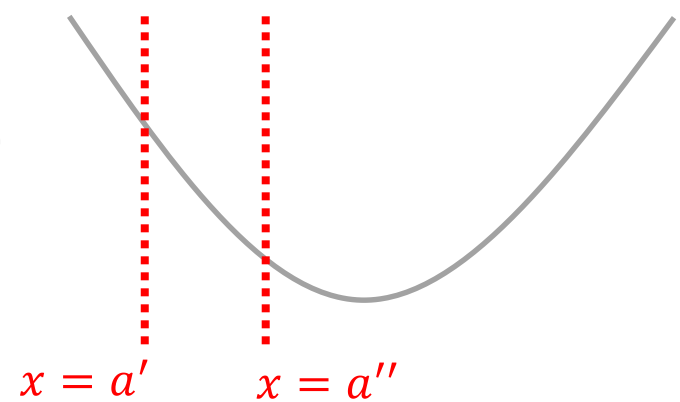
●ケース１

$$\frac{df(\bar{x})}{dx} > 0, \frac{dg_1(\bar{x})}{dx} = 1 > 0$$




●ケース２

$$\frac{df(\bar{x})}{dx} < 0, \frac{dg_2(\bar{x})}{dx} = -1 < 0$$



# 制約条件有り④
# sheffiの記述
# “same sign” condition

・最適解における「目的関数の一階微分値」と「有効制約関数の一階微分値」の符号は必ず等しい。つまり…

$$\frac{df(\bar{x})}{dx} = \sum_j u_j \frac{dg_i(\bar{x})}{dx}$$

$$u_j \geq 0$$

$$u_j g_i(\bar{x}) = 0$$

$$g_i(\bar{x}) \geq 0$$

（※$j$は有効制約関数の添え字集合）

（有効制約関数＝それ上に最適解があるもの）

→考えが雑すぎないか？

# 制約条件有り⑤

・少し寄り道になりますが、

「制約条件がある」
＝「実行可能領域（=feasible region)が限定されている」

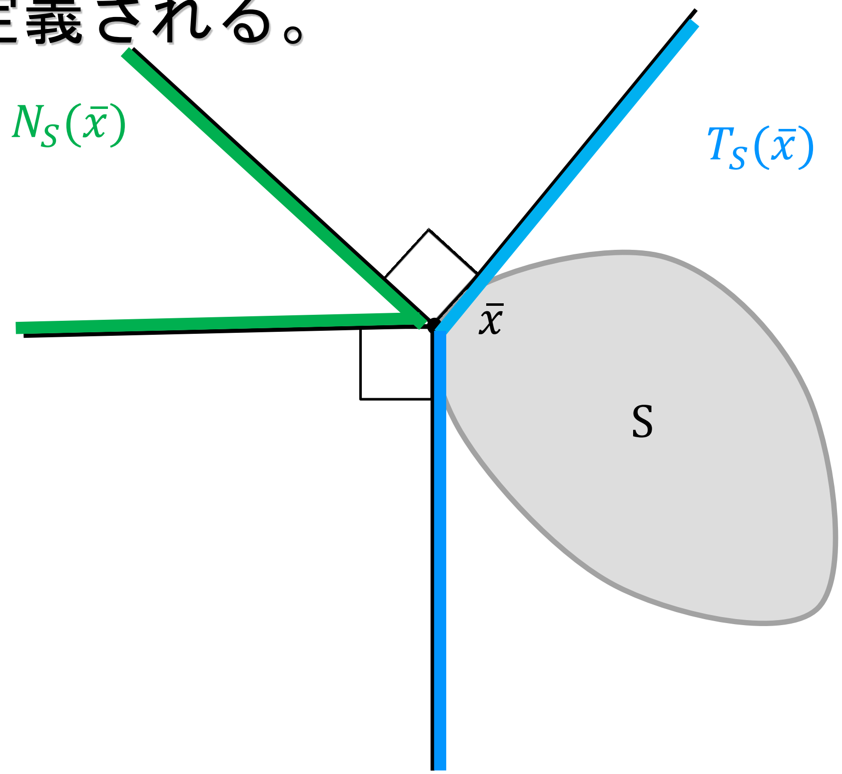
場合の最適解の持つ性質を見てみます。

Part 1
Part 2
→ Part 3
Part 4
Part 5

# 制約条件有り⑥
# 最適化のイメージと定義

接錐(tangent cone)$T_S(\bar{x})$ …集合を点$\bar{x}$において線形近似したもの。

法線錐(normal cone) $N_S(\bar{x})$…接錐に対し以下のように幾何学的に定義される。

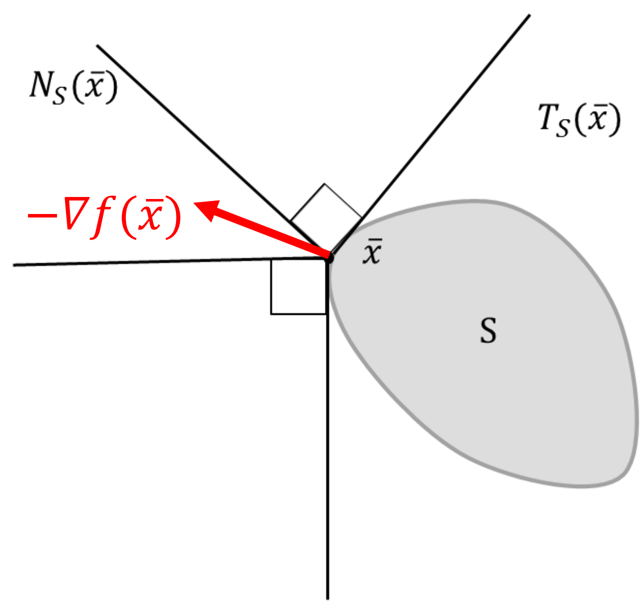


# 制約条件有り⑦
# 最適化のイメージと定義

$S$は実行可能領域で、空でない凸集合。$f$が点$\bar{x} \subseteq S$において微分可能な凸集合であるとする。このとき、

$$-\nabla f(\bar{x}) \in N_S(\bar{x})$$

は、$\bar{x}$が大域的な最適解であるための必要十分条件である。



# 制約条件有り⑧
# 最適化のイメージと定義

<軽い証明>

$\bar{x}$が最適解であれば、ごく近傍の点$x_k$では、

$$f(x_k) - f(\bar{x}) \geq 0$$

$$f(x_k) - f(\bar{x}) = \langle \nabla f(\bar{x}), x_k - \bar{x} \rangle$$

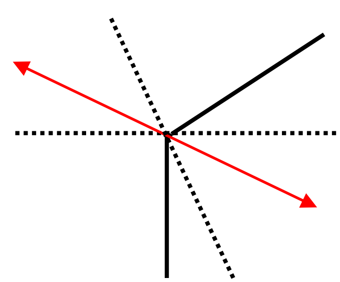
Taylor展開
２次以降無視

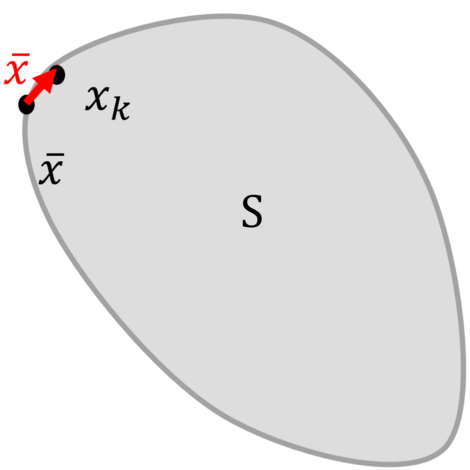
より、$\langle \nabla f(\bar{x}), x_k - \bar{x} \rangle \geq 0$が言えます。

ゆえに、$\langle -\nabla f(\bar{x}), x_k - \bar{x} \rangle \leq 0$となり、

$x_k - \bar{x}$は右のように$T_S(\bar{x})$内の接ベクトルを表しているといえるので、

$-\nabla f(\bar{x}) \in N_S(\bar{x})$が理解できます。





Part 1

Part 2

→ **Part 3**

Part 4

Part 5

# 制約条件有り⑨
# 最適化のイメージと定義

このように、最適性条件とは、

「目的関数の勾配ベクトル」と「実行可能領域」の幾何的関係性

を表しているものです。

実行可能領域は現実には制約条件式によって定められるので、より実用的には

「目的関数の勾配ベクトル」と「制約関数の勾配ベクトル」の幾何的関係性

を表しているものだと推測できます。

Part 1
Part 2
→ **Part 3**
Part 4
Part 5

# 制約条件有り⑩
# ＫＫＴ条件
# （Karush-Kuhn-Tucker)

$$\nabla f(\bar{x}) = \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x})$$
$$\bar{\lambda}_i \geq 0$$
$$g_i(\bar{x}) \geq 0$$
$$\bar{\lambda}_i g_i(\bar{x}) = 0$$
$$(i = 1, \ldots, m)$$

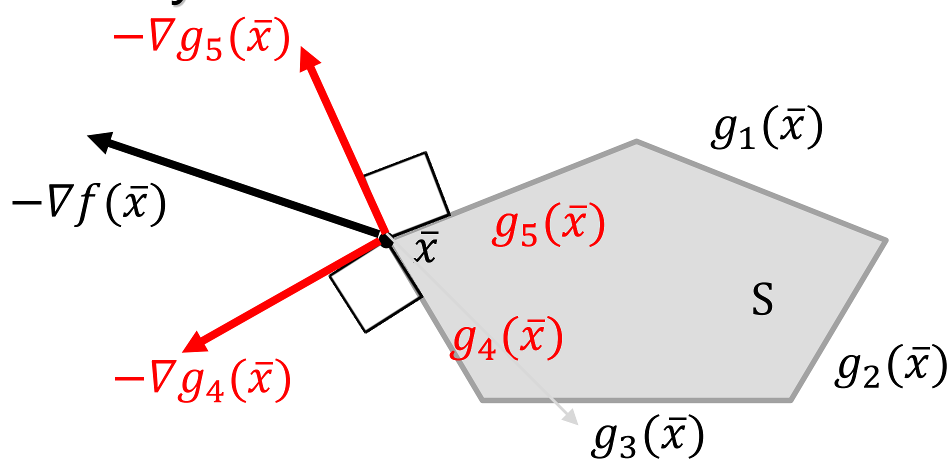
局所的最適解$\bar{x}$に対して、次式を満たすLagrange乗数$\bar{\lambda}$が存在する。

Part 1
Part 2
→ Part 3
Part 4
Part 5

# 制約条件有り⑪
# ＫＫＴ条件

KKT条件は、局所的最適解$\bar{x}$における目的関数の勾配$\nabla f(\bar{x})$に-1をかけたベクトルが**有効制約条件に対応する制約関数の勾配に−1をかけた$-\nabla g_i(\bar{x})$**によって張られる凸多面錐に含まれることを意味する。

特に条件$\bar{\lambda}_i g_i(\bar{x}) = 0$は、有効でない制約条件に対応するLagrange乗数$\bar{\lambda}_i$が0になることを示している。（=相補性条件、complementarity condition）



Part 1
Part 2
Part 3
→ Part 4
Part 5

# 均衡配分への適用のために①

- 非負制約（Nonnegativity constraints）
- 線形等式制約（Linear Equality Constraints）

$$\min.\ f(\vec{x})$$

$$subject\ to$$

$$x_i \geq 0$$ 非負制約

$$\sum_i h_{ij} x_i = b_j$$ 線形等式制約

このニつの不等式の組み合わせで考える。

$$\sum_i h_{ij} x_i \geq b_j$$

$$-\sum_i h_{ij} x_i \geq -b_j$$

# 均衡配分への適用のために②

KKT条件

$$\nabla f(\bar{x}) = \sum_{j=1}^{m} \bar{\lambda}_j \nabla g_j(\bar{x}) \quad (①)$$

$\bar{\lambda}_j \geq 0\ (②),\ g_j(\bar{x}) \geq 0\ (③),\ \bar{\lambda}_j g_j(\bar{x}) = 0(④),$
を適用すると、

$x_i \geq 0, \sum_j h_{ij} x_i \geq b_j, -\sum_j h_{ij} x_i \geq -b_j$
（ $i = 1, \ldots, n$ j=1,…,m)

①→ $\nabla f_i(\bar{x}) = \bar{\lambda}_i \frac{dx_i}{dx_i} + \sum_{j=1}^{m} \bar{\lambda}'_j h_{ij} - \sum_{j=1}^{m} \bar{\lambda}''_j h_{ij}\ (\forall i)$

$= \bar{\lambda}_i + \sum_{j=1}^{m} \overline{\lambda'''}_j h_{ij}\ (\forall i)$ より、　← $\overline{\lambda'''}_j$は非負制約なし！

$\boxed{\begin{array}{l} ④\rightarrow(\nabla f_i(\bar{x}) - \sum_{j=1}^{m} \overline{\lambda'''}_j h_{ij})\ \bar{x}_i = 0\ (\forall i) \\ ②\rightarrow\nabla f_i(\bar{x}) - \sum_{j=1}^{m} \overline{\lambda'''}_j h_{ij} \geq 0\ (\forall i) \\ ③\rightarrow x_i \geq 0\quad (\forall i) \\ ③、④\rightarrow\sum_j h_{ij} x_i = b_j\ (\forall j) \end{array}}$

←均衡配分で用いそうな
KKT条件の特殊型

(※$\nabla f_i(\bar{x})$は、停留点における$\nabla f(\bar{x})$（ベクトル）のi番目の要素を表す。）

Part 1

Part 2

Part 3

→ **Part 4**

Part 5

# 均衡配分への適用のために③ Lagrangian関数

元の問題

min. $f(\vec{x})$ subject to $g_n(\vec{x}) \geq 0 \ (\forall n)$ のとき、

Lagrangian関数は

$L(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_n \lambda_n g_n(\vec{x}) \quad (\lambda_n \geq 0)$

である。

$L$の、$\vec{x}, \vec{\lambda}$の両方についての一次の最適性条件は、元の最適化問題に対するKKT条件と同値になる。

（容易に示せます。）

Part 1

Part 2

Part 3

→ **Part 4**

Part 5

# 均衡配分への適用のために④ Lagrangian関数

$L(\overrightarrow{x^*}, \overrightarrow{\lambda^*})$は$L(\vec{x}, \vec{\lambda})$の最大値でも最小値でもない。

$L(\overrightarrow{x^*}, \overrightarrow{\lambda^*})$は鞍点で、

$L(\overrightarrow{x^*}, \vec{\lambda}) \leq L(\overrightarrow{x^*}, \overrightarrow{\lambda^*}) \leq L(\vec{x}, \overrightarrow{\lambda^*})$　を満たす。

Lagrangian的アプローチの利点は、「制約有り最小化問題」を「制約無しの鞍点探索問題」に読み替えることができることにある。

→鞍点がなぜ最適解を示す？

Part 1

Part 2

Part 3

→ **Part 4**

Part 5

# 均衡配分への適用のために⑤ Lagrangian関数

KKT条件と同値にするために、 $L$は$\vec{\lambda}$について最大化する必要があるのです。

$\vec{\lambda}$について最適

→最大化なら$\dfrac{\partial L\left(\overrightarrow{x^*}, \overrightarrow{\lambda^*}\right)}{\partial \lambda_n} \leq 0$最小化なら$\dfrac{\partial L\left(\overrightarrow{x^*}, \overrightarrow{\lambda^*}\right)}{\partial \lambda_n} \geq 0$

$\dfrac{\partial L\left(\overrightarrow{x^*}, \overrightarrow{\lambda^*}\right)}{\partial \lambda_n} = -g(\overrightarrow{x^*})$ ≤0 (∵元の制約条件が$g(\overrightarrow{x^*}) \geq 0$)

よって$\vec{\lambda}$については最大化をすることになる■

# さいごに

・上記はすべて最小値に対する議論でしたが、最大値を考える場合は目的関数を－１倍することで最小化問題に落としこめます。本質的にも、凸関数に負値を掛けると凹関数になって、すべて同様の議論が可能なことがわかります。

・この本で扱う目的関数のヘッセ行列はだいたい対角行列。ゆえに、対角要素がすべて正であれさえすれば、ヘッセ行列は正定値になります。

Part 1

Part 2

Part 3

Part 4

→ **Part 5**

# まとめ

・最適解を得るための一次と二次の最適性条件を、
・制約無しから制約有りに拡張し、
・一次元問題から多次元問題に拡張し、
・最適解のイメージをえつつ、
・理解することができた。
・また、配分問題に向けた非線形問題の特殊な場合に対する考察も行った。
・最後に制約条件の消し方も学んだ。

# 参考文献

■交通ネットワークの均衡分析
（土木学会、平成１１年）
Appendixに、sheffiの2章そのまま和訳要約したようなものが載ってます。笑
■非線形最適化の基礎
（福島雅夫、朝倉出版、２００１年）
研究室においてある本の中では一番ちゃんと書かれています。
■東京大学工学教程　基礎系数学　最適化と変分法
（寒野義博・土谷隆、東京大学工学教程編纂委員会、２０１３年）
わかりやすいです。

# おつかれさまでした！

後日の質問は　shoji@trip.t.u-tokyo.ac.jp にどうぞ。