2020年度「数値最適化」

第三部「制約のある一般の最適化問題」

制約のある一般の最適化問題

第三部では制約のある一般の最適化問題を扱います。すなわち

$$\begin{aligned} \text{min.} \quad & f(x_1,\ldots,x_n), \\ \text{s.t.} \quad & g_1(x_1,\ldots,x_n)=0,\ldots,g_m(x_1,\ldots,x_n)=0, \\ & h_1(x_1,\ldots,x_n)\geqq 0,\ldots,h_l(x_1,\ldots,x_n)\geqq 0. \end{aligned}$$

で与えられる問題において条件を満たし、目的関数を最小化する変数の値を見つける方法を考えます。

第二部の冒頭で述べた前提は第三部でも有効です。

すなわち、目的関数f, 制約関数g_j, h_kは1階または2階微分可能な連続関数とします。

制約のある一般の最適化問題（1変数の場合）

まずは、制約の無い問題と同様に1変数の場合を考えましょう。

$$\min. \quad f(x), \quad \text{s.t.} \qquad g_1(x)=0,\ldots,g_m(x)=0, \quad h_1(x) \geqq 0,\ldots,h_l(x) \geqq 0.$$

で与えられる問題において条件を満たし、目的関数を最小化する変数の値を見つける方法を考えます。

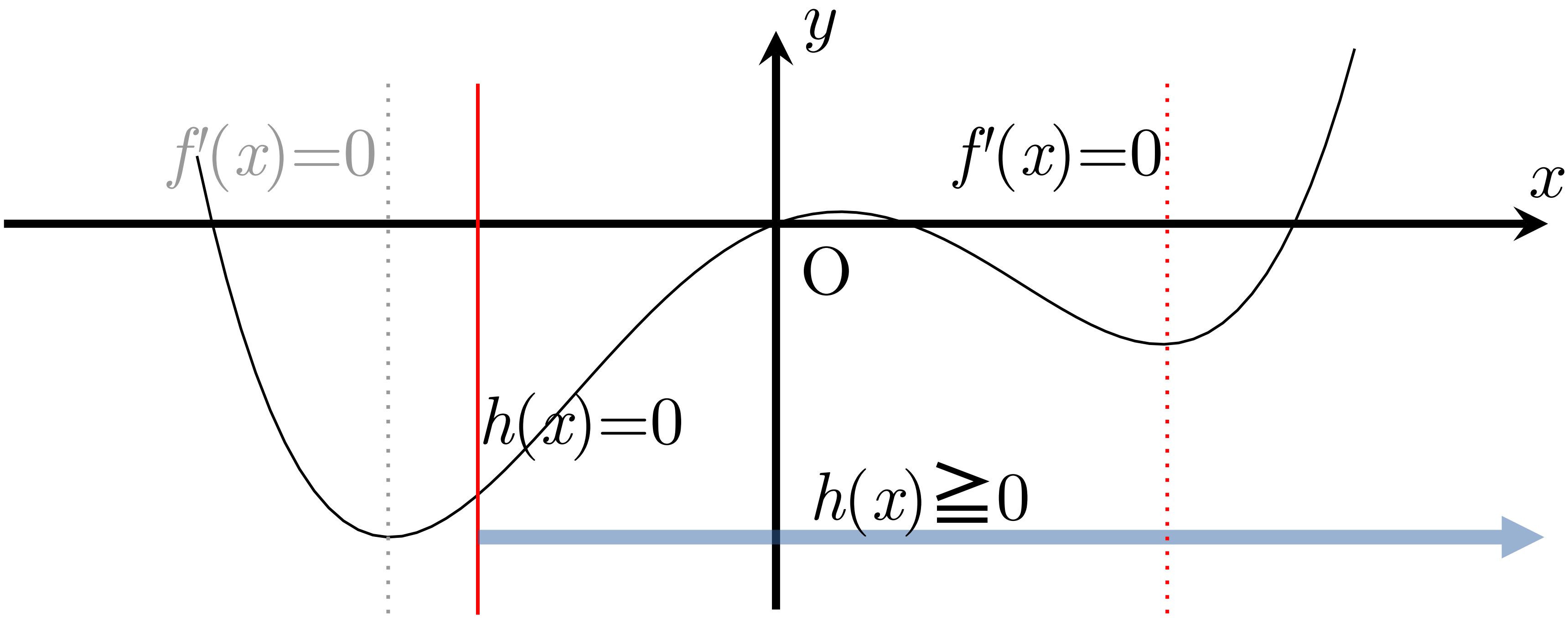
制約関数は滑かな連続関数なので制約を満たす変数の集合（可能領域）は（閉）区間または境界になります。

色々な想定はできますが、特殊な場合は除いて、$g_j(x)=0$ で境界、$h_k(x) \geqq 0$で（片）区間が得られるものとして説明を続けます。

さらに、1変数では（境界の得られる）等式成約はただの方程式なので、省きます。

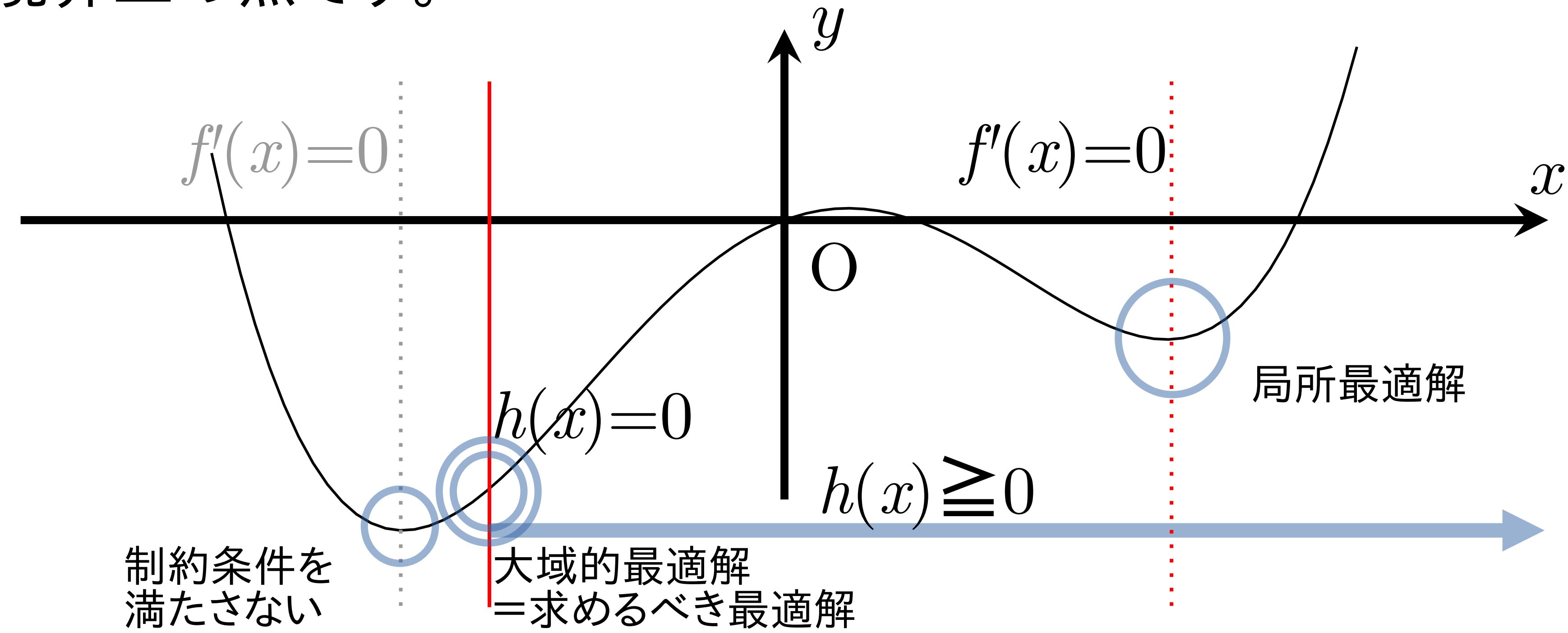
制約のある一般の最適化問題（1変数の場合）

第二部で使った多項式のグラフを流用すると、以下のように問題を図示できます。



制約のある一般の最適化問題（1変数の場合）

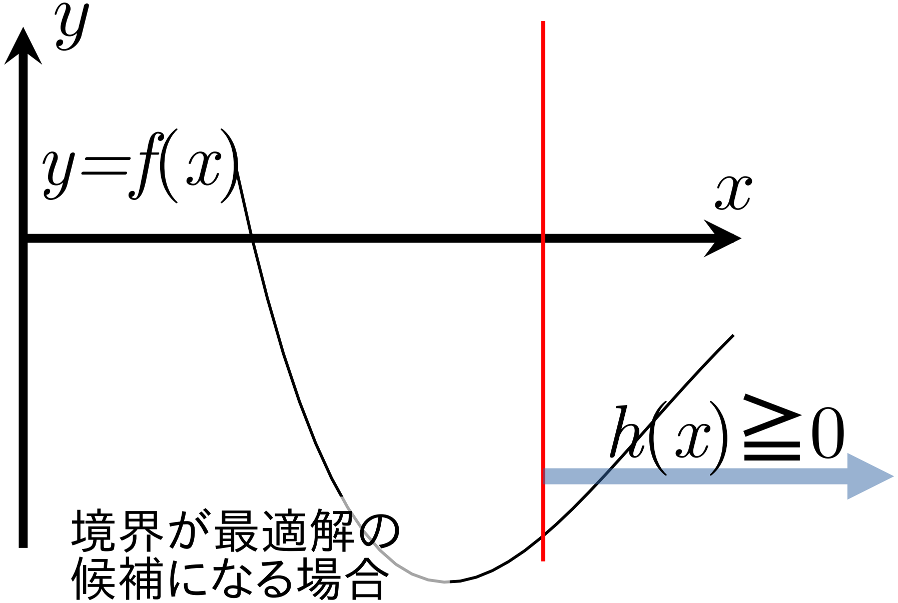
下図の場合、2つの停留点のうち右側は局所最適解ですが、左側は制約条件を満たしません。さらに大域的最適解は停留点ではない境界上の点です。

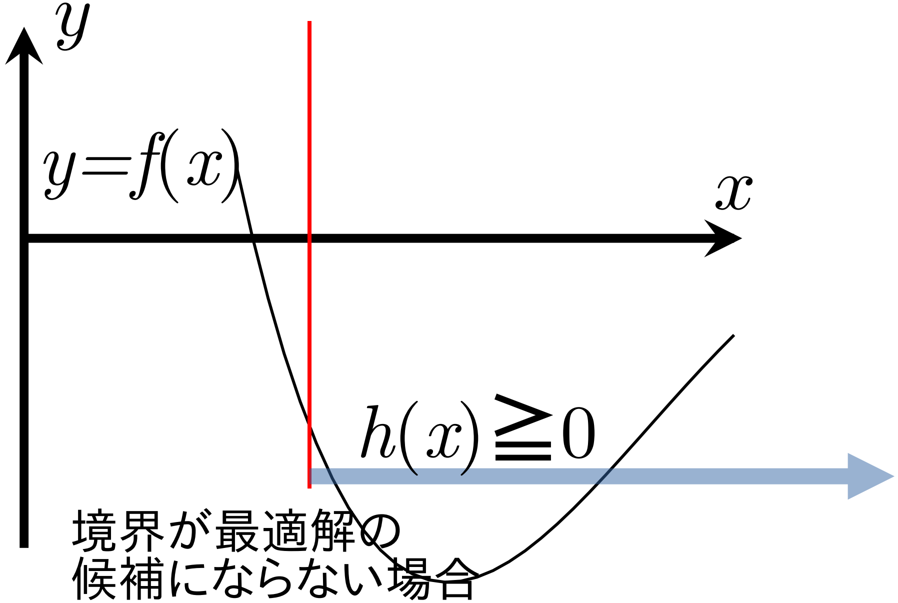


制約のある一般の最適化問題（1変数の場合）

図示された例からも明らかな通り、制約のある問題では停留点も最適解の候補でかつ、境界に最適解がある場合も考えられます。

また、容易に想像できる通り、境界における最適解の候補は目的関数の勾配ベクトル、すなわち降下方向とも関係があります。





制約のある一般の最適化問題

以上の考察から、制約のある一般の最適化問題の最適解を以下の手順で求めることが考えられます。

1 実行可能領域（の内点）で停留点を探索する

2 境界上で最適解の候補となる実行可能解を探索する

3 最適解の候補から大域的最適解を決める

1における「内点」は線形計画問題の場合と同様に不等式制約$h_1 \geqq 0, \ldots,$から等号を省いた条件$h_1 > 0, \ldots,$を満たす点の集合です。

これには定義域全体から停留点を見つけ、そのうち制約条件を満たすものを決めるという手順を使うこともできる。

であれば、制約の無い場合のNewton法等をこれに利用できます。

制約のある一般の最適化問題

内点での探索は制約の無い場合と同一になるので、境界上の最適解候補探索についで考えることにしましょう。

ここで言う境界は不等式制約の一部の条件を等号の成立のもとで満たす場合です。すなわち$h_k=0$のような等式が成立することになります。

これは、等式制約$g_j=0$を考えることと一緒になります。

そこで、まず等式制約だけからなる問題について考えます。

ただ1つの等式制約のある最適化問題

簡単のために等式制約が1つだけある場合を考えます。

$$\min.\quad f(x_1,\ldots,x_n),\quad \text{s.t.}\quad g(x_1,\ldots,x_n)=0.$$

以下ではこれをベクトルを使って次のように表現します。

$$\min.\quad f(\boldsymbol{x}),\quad \text{s.t.}\quad g(\boldsymbol{x})=0.$$

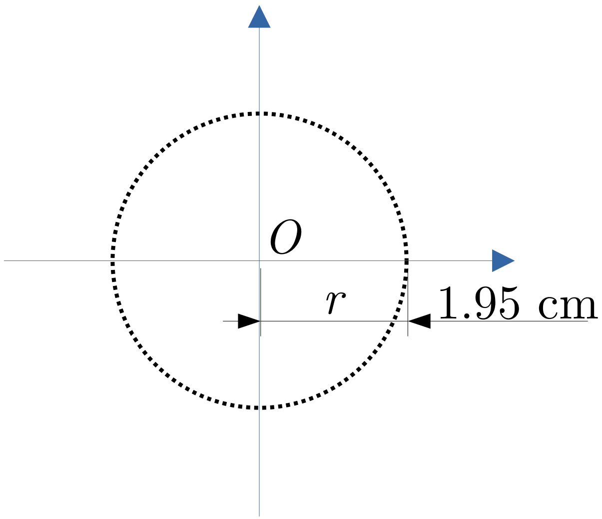
このとき$g(\boldsymbol{x})=0$を満たす実行可能領域は関数gの関数値0の等高線と一致します。

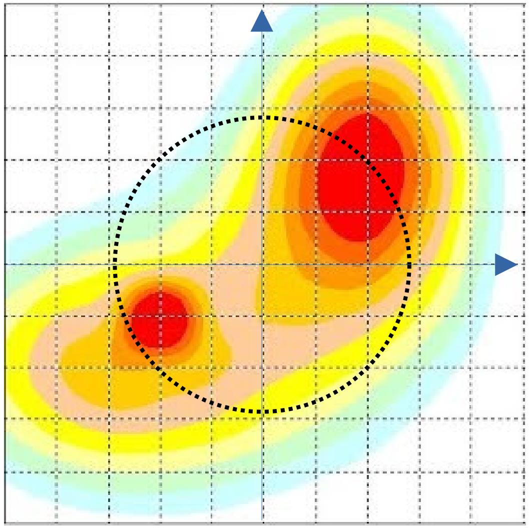
$n=2$となる2変数の場合であれば、関数値を高度と見做して、ここで言う等高線と地形図の等高線は同じものと言えます。

ただ1つの等式制約のある最適化問題

$g(x_1,\ x_2) = {x_1}^2 + {x_2}^2 - r^2$の場合を図示します。制約式$g=0$を満たす点の集合は原点を中心とする半径$r$の円になります。

ここにある目的関数（のヒートマップ）を重ねて描いたものを見て、最適解の候補となる点を見つけてください。

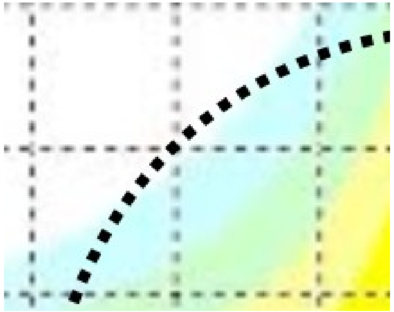
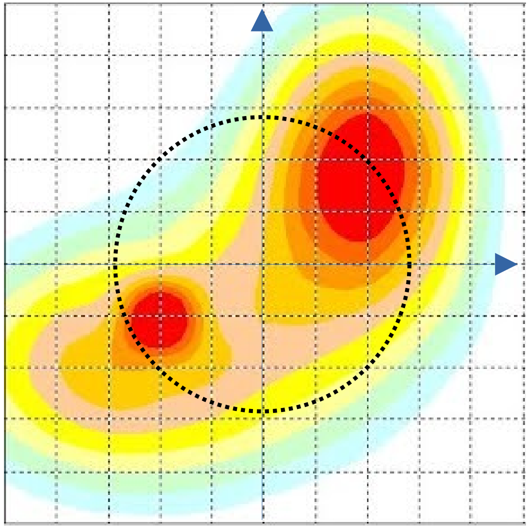
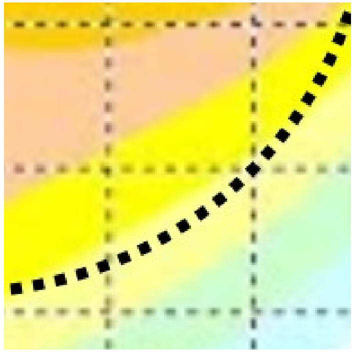




ただ1つの等式制約のある最適化問題

g=0の円周上で目的関数が最小になる点が最適解なので、左上のあたりが最適解の候補になるのが分かります。右下には局所最適解になる点もあるかもしれません。

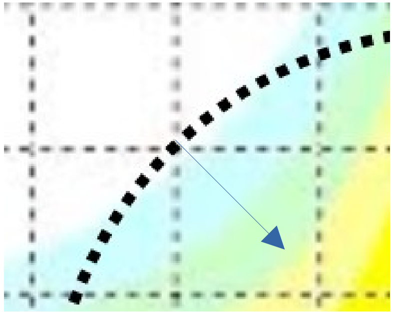
最適解近辺の円周と目的関数を比較してどのような状況であるか確かめてください。

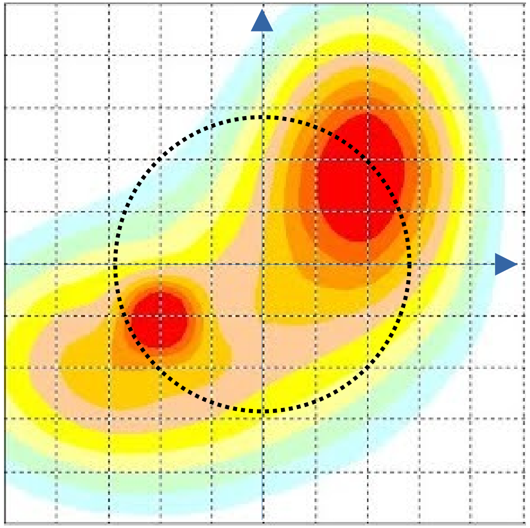




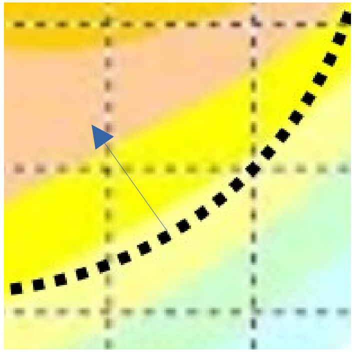
ただ1つの等式制約のある最適化問題

あくまで図示なので確認はし辛いのですが、最適解の候補となる「点」では目的関数の等高線と円周が直交している様子が確認できないでしょうか。

別の表現をすれば目的関数の増加方向ベクトル（勾配ベクトル）が円周に直交している、ということになります。





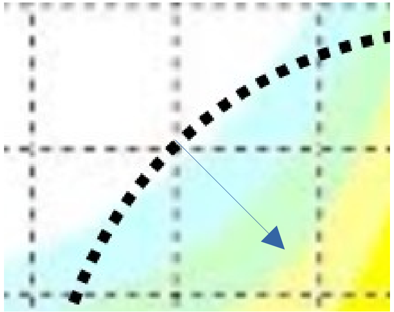


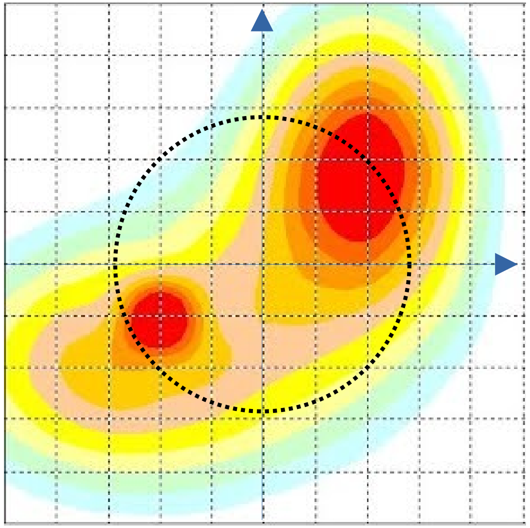
ただ1つの等式制約のある最適化問題

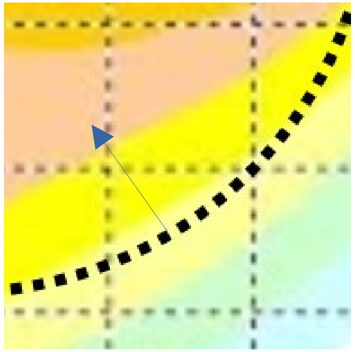
制約条件を満たす点の集合と目的関数の等高線図とから示唆される最適解の候補となる点の条件は

$$\nabla f \mathbin{/\!/} \nabla g$$

すなわちfとgの勾配ベクトルが平行であることです。



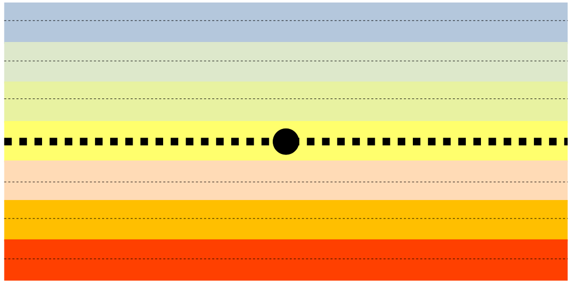




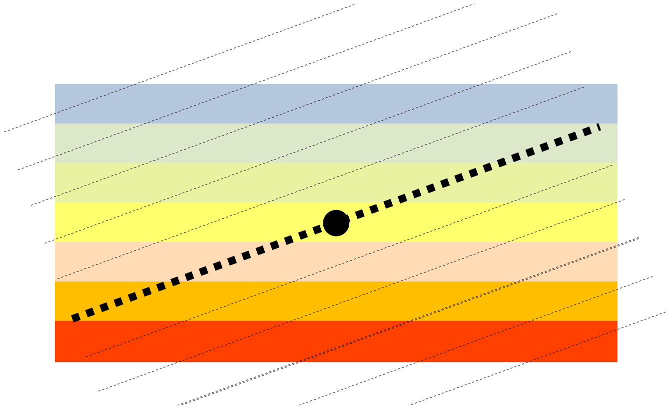
ただ1つの等式制約のある最適化問題

円周上の点の近傍でfとgの等高線を考えれば、勾配ベクトルによる条件はより明確になります。

勾配ベクトルが平行でない点近傍では等高線は交差し、可能領域中の移動で目的関数が増減するので最適解とはなりません。



勾配ベクトルが平行な場合
可能領域を移動(←→)しても目的関数は増減しない



勾配ベクトルが平行でない場合
可能領域を移動(←→)すると目的関数が増減する

ただ1つの等式制約のある最適化問題

ただ1つの等式制約$g(\boldsymbol{x})=0$のもとでの最適解の候補探索は次の条件式を満たす解$\boldsymbol{x}$の探索となります。（$\mathbf{0}$はゼロベクトル）

$$\lambda \neq 0,\ \ \nabla f(\boldsymbol{x}) - \lambda \nabla g(\boldsymbol{x}) = \mathbf{0}$$

定数λも変数に含めて考えれば関数のゼロ点探索になります。

このことを利用した「低次元すなわちnが比較的小さい場合に有効な方法」と言われるLagrangeの未定係数法を以降のスライドで説明します。

Lagrangeの未定係数法（等式制約が1つある場合）

例題：次の最小化問題の最適解・最適値を求めよ

$$\text{min.}\quad f(x_1,x_2)=x_1{}^2+x_2{}^2,$$

$$\text{s.t.}\quad g(x_1,x_2)=ax_1+bx_2+c=0.$$

この問題にLagrangeの未定係数法を適用します。

最適解が原点から直線（$ax_1+bx_2+c=0$）に下ろした垂線の足であることは明らかなので、これを確かめるということになります。

（連立）方程式$\nabla f(x_1,x_2)-\lambda\nabla g(x_1,x_2)=\mathbf{0}$を満たすゼロ点を求めますが、制約条件$g(x_1,x_2)=0$も満たす必要があります。

これらの連立方程式を次のスライドに示します。

Lagrangeの未定係数法（等式制約が1つある場合）

例題：次の最小化問題の最適解・最適値を求めよ

$$\begin{aligned} &\text{min.} && f(x_1,x_2)=x_1{}^2+x_2{}^2, \\ &\text{s.t.} && g(x_1,x_2)=ax_1+bx_2+c=0. \end{aligned}$$

この問題にLagrangeの未定係数法を適用します。

最適解が原点から直線（$ax_1+bx_2+c=0$）に下ろした垂線の足であることは明らかなので、これを確かめるということになります。

（連立）方程式$\nabla f(x_1,x_2)-\lambda\nabla g(x_1,x_2)=\mathbf{0}$を満たすゼロ点を求めますが、制約条件$g(x_1,x_2)=0$も満たす必要があります。

これらの連立方程式を次のスライドに示します。

Lagrangeの未定係数法（等式制約が1つある場合）

$$
\begin{aligned}
\frac{\partial}{\partial x_1} f(x_1, x_2) - \lambda \frac{\partial}{\partial x_1} g(x_1, x_2) &= 2x_1 - \lambda a &= 0, \\
\frac{\partial}{\partial x_2} f(x_1, x_2) - \lambda \frac{\partial}{\partial x_2} g(x_1, x_2) &= 2x_2 - \lambda b &= 0, \\
g(x_1, x_2) &= ax_1 + bx_2 + c &= 0.
\end{aligned}
$$

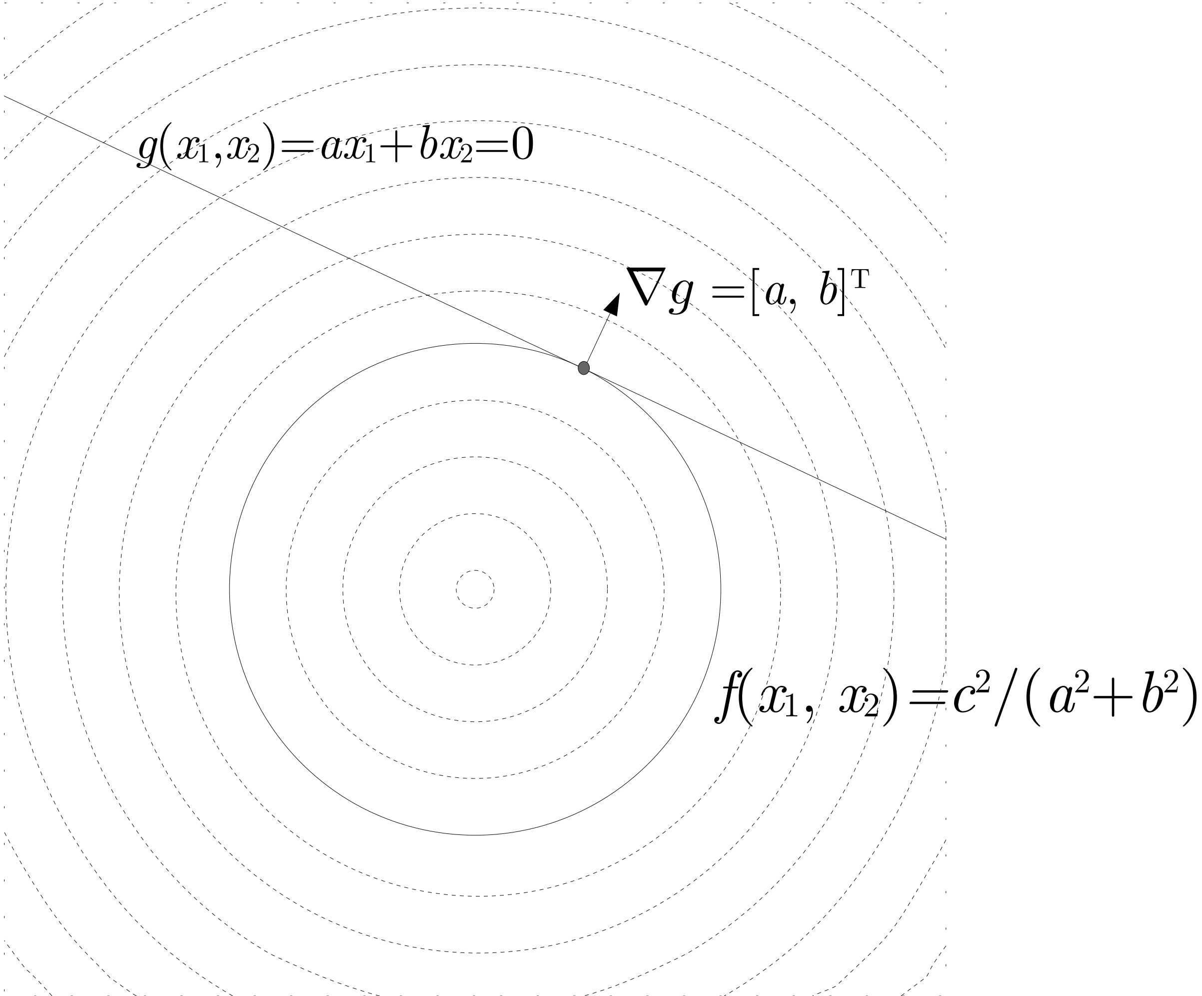
この場合、連立方程式は線形で　$a^2+b^2 \neq 0$ であれば唯一つの解

$$x_1 = -ac/(a^2+b^2),\ x_2 = -bc/(a^2+b^2),\ \lambda = -2c/(a^2+b^2)$$

を持つ。

これが$f(\boldsymbol{x})$を最小化する$g(\boldsymbol{x})=0$上の点であることは明らか。

Lagrangeの未定係数法（等式制約が1つある場合）



複数の等式制約のある最適化問題

次に等式制約が複数ある場合、つまり一般の最適化問題のうち、不等式制約が無い場合を考えます。

$$\min. \quad f(\boldsymbol{x}), \quad \text{s.t.} \qquad g_1(\boldsymbol{x})=0,\ldots,g_m(\boldsymbol{x})=0.$$

目的関数fとm個のうちの1つの制約関数g_jとの勾配ベクトルが平行でなければ両関数の等高線が交差することは$m=1$の場合と同じです。したがって、最適解の候補が満たす連立方程式は次のようになります。（全てを満たす$\boldsymbol{x}$, $\lambda_1,\ldots,\lambda_m\neq 0$が必要なことに注意）

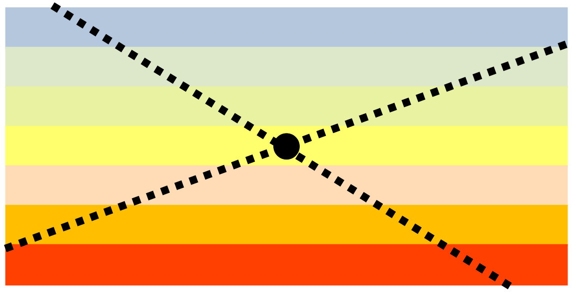
$$\nabla f(\widetilde{\boldsymbol{x}}) - \sum_{i=1}^{m} \lambda_m \nabla g_m(\widetilde{\boldsymbol{x}}) = 0,$$

$$g_i(\widetilde{\boldsymbol{x}}) = 0 \quad i = 1,\ldots,m.$$

複数の等式制約のある最適化問題

注意: 2変数の場合を考えると前頁の連立方程式を満たさない、下図のような場合も最適解であるように思われるかもしれません。

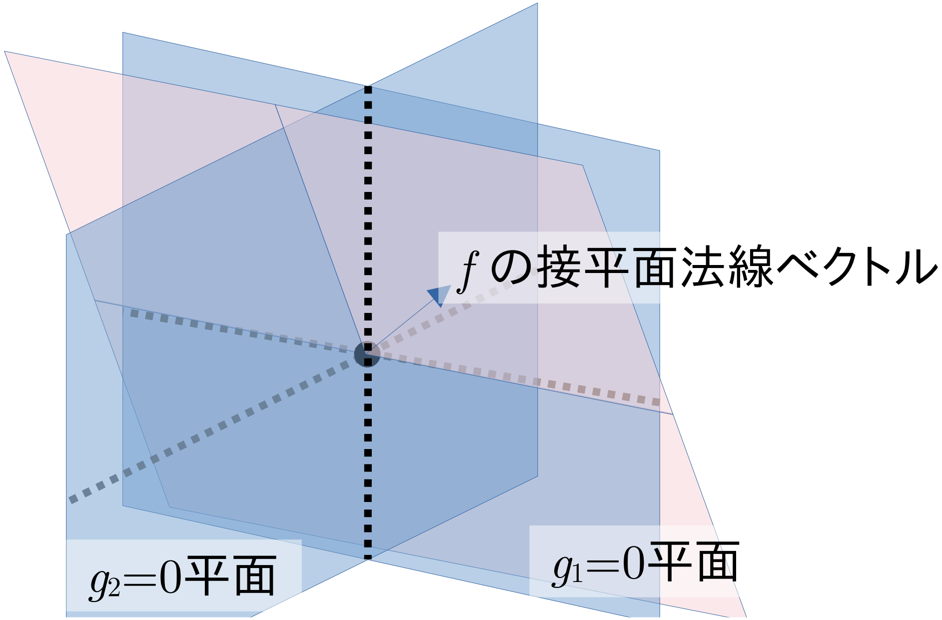
これが実際に2変数の問題であれば実行可能領域が1点に退化した状態であるので、その中から最適解を見つけるという問題を考える必要はなくなってしまいます。



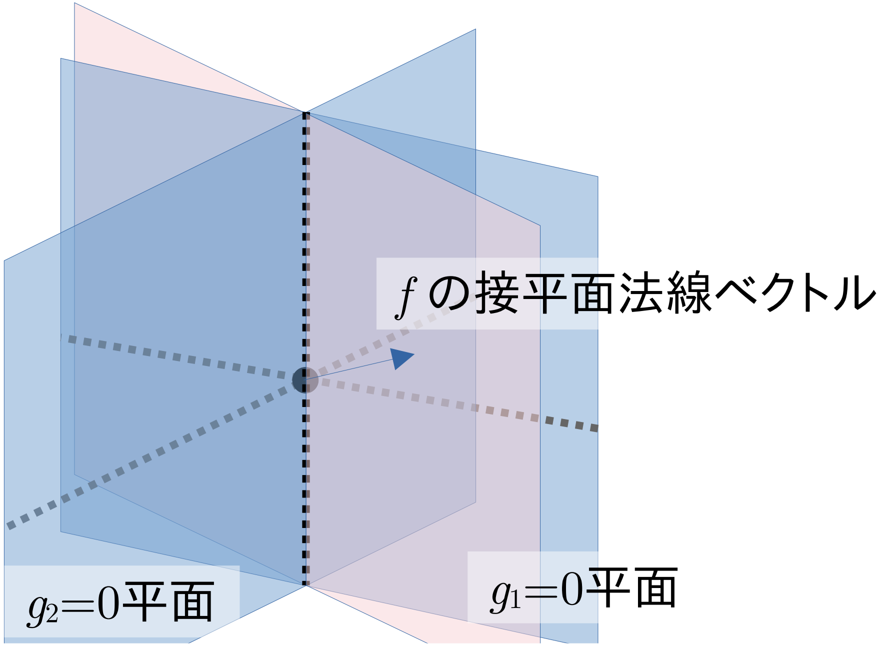
2つの等式制約$g_1=0$, $g_2=0$の交わりに最適解がある例?

複数の等式制約のある最適化問題

（前頁からの続き）一方で、複数の等式制約の交わりにおいて下図のような場合は連立方程式を考える必要があることが容易に想像できるはずです。



$g_1=0$, $g_2=0$の交わりがfの接平面に交差しているので、交わり上でfが増減する。



$\lambda_1\nabla g_1+$, $\lambda_2\nabla g_2 /\!/ \nabla f$を満たす$\lambda_1,\lambda_2$が存在し、交わり上に停留点が存在する。

複数の等式制約のある最適化問題

形式的ではありますが、次のようなベクトル関数を考えれば

$$
F(\widetilde{\boldsymbol{x}},\lambda) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(\widetilde{\boldsymbol{x}}) - \displaystyle\sum_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_1}(\widetilde{\boldsymbol{x}}), \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(\widetilde{\boldsymbol{x}}) - \displaystyle\sum_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_n}(\widetilde{\boldsymbol{x}}), \\ g_1(\widetilde{\boldsymbol{x}}) \\ \vdots \\ g_m(\widetilde{\boldsymbol{x}}) \end{bmatrix}
$$

最適化問題全体を関数のゼロ点探索と見なし、いわゆるニュートン法や準ニュートン法を使うことができます。

複数の等式制約のある最適化問題

Lagrangeの未定係数法は使えない？

例題で導いた連立1次方程式はその条件下での関数$\boldsymbol{F}(\boldsymbol{x},\boldsymbol{\lambda})$のゼロ点探索の方程式です。

たまたま連立1次方程式になってしまい、解くことができました。

一般的に言えばLagrangeの未定係数方法は「方程式$\boldsymbol{F}(\boldsymbol{x},\boldsymbol{\lambda})=0$を解け」と表現できるものです。多くの場合は非線形方程式の反復解法を使うことになります。

容易に解けるのは低次元（nやmが小さい場合）に限られるということです。

ただ1つの不等式制約のある最適化問題

今度は不等式制約が1つだけある場合を考えます。

$$\min.\quad f(\boldsymbol{x}),\quad \text{s.t.}\qquad h(\boldsymbol{x}) \geqq 0.$$

不等式制約の実行可能領域は不等式$h(\boldsymbol{x})>0$または等式$h(\boldsymbol{x})=0$のいずれかを満たす点の集合です。

そこで、2つの条件を分けて考えることにします。

まず、内点にあたる不等式を満たす$\boldsymbol{x}$ について、1変数の場合について述べた場合と同様に、これは制約の無い最適化問題の場合と同様に、停留点探索を実行することになります。

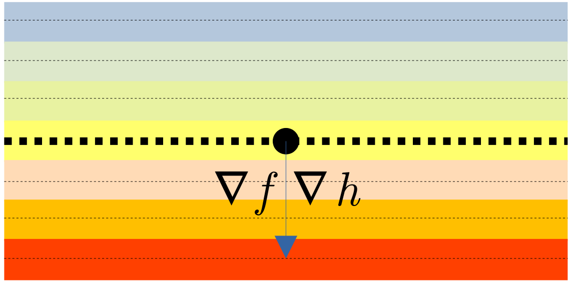
不等式$h(\boldsymbol{x})>0$を満たす停留点は局所最適解となりますし大域的最適解の候補となります。

ただ1つの不等式制約のある最適化問題

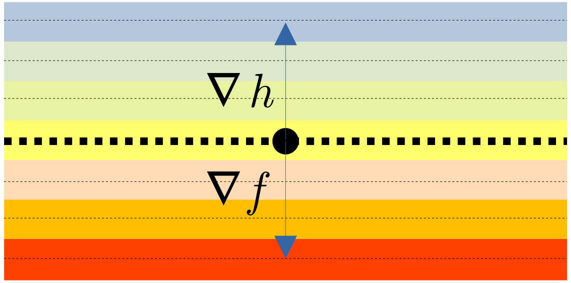
次に等式$h(\boldsymbol{x})=0$を満たす$\boldsymbol{x}$ について考えます。

これは、等式制約$g(\boldsymbol{x})$の場合と同様に考えてベクトル関数$\boldsymbol{F}$のゼロ点探索を実行することができます。

ただし、見つかったゼロ点について、下図のような場合分けを考える必要があります。



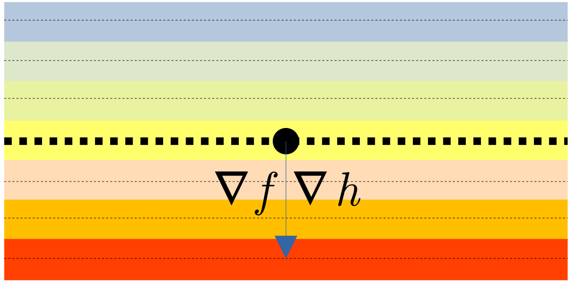
勾配ベクトル∇fと∇hが同一方向の場合、内点$h>0$はfの増加する方向となる。



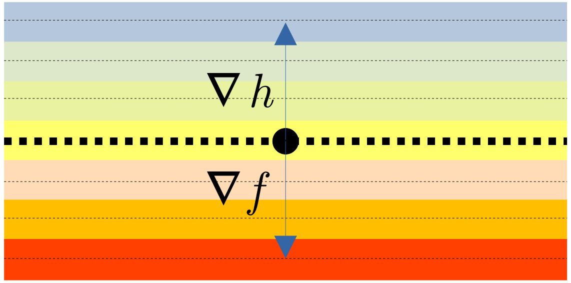
勾配ベクトル∇fと∇gが反対方向の場合、内点$h>0$はfの減少する方向となる。

ただ1つの不等式制約のある最適化問題

$h(\boldsymbol{x})=0$上の最適解候補において、勾配ベクトルの向きが右図のように目的関数と制約関数で逆向きである場合、$h(\boldsymbol{x})>0$の内点領域に、より小さい目的関数を実現する実行可能領域があるので、最適解とはなり得ません。



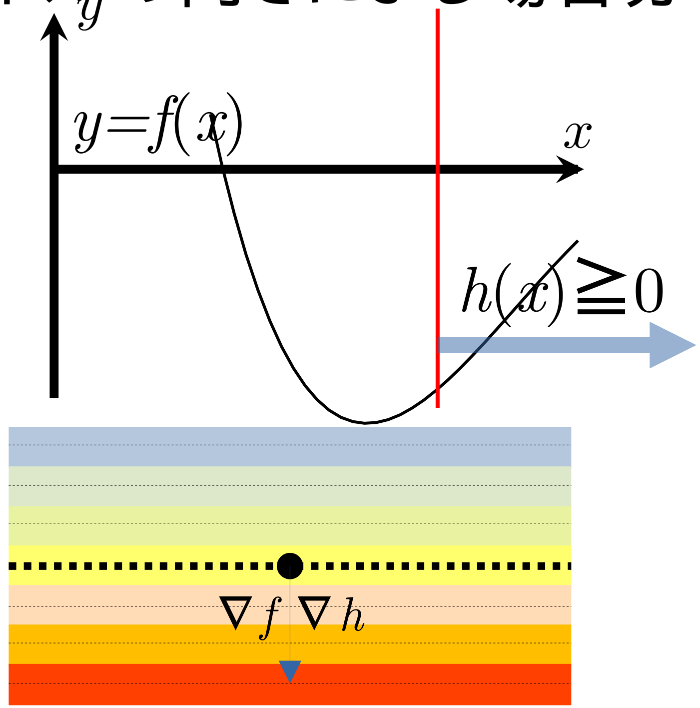
勾配ベクトル∇fと∇hが同一方向の場合、内点$h>0$はfの増加する方向となる。



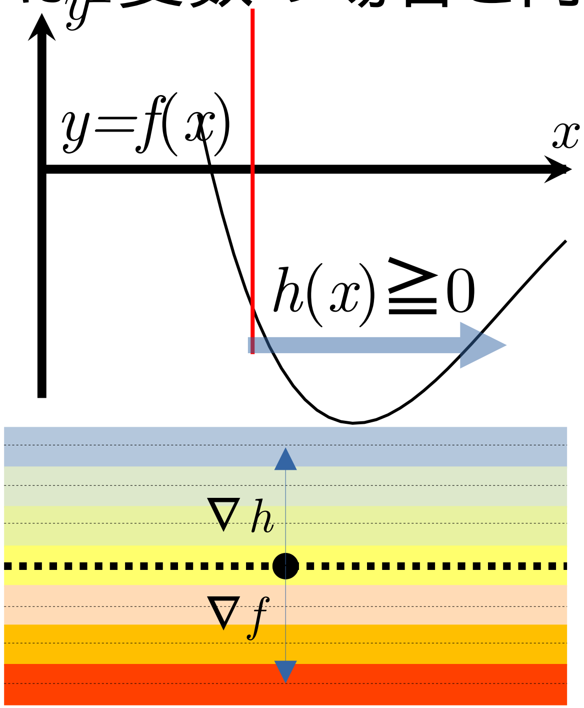
勾配ベクトル∇fと∇gが反対方向の場合、内点$h>0$はfの減少する方向となる。

ただ1つの不等式制約のある最適化問題

勾配ベクトルの向きによる場合分けは1変数の場合と同じものと言えます。



勾配ベクトル∇fと∇hが同一方向の場合、内点$h>0$はfの増加する方向となる。



勾配ベクトル∇fと∇gが反対方向の場合、内点$h>0$はfの減少する方向となる。

ただ1つの不等式制約のある最適化問題

fとhの勾配ベクトルが同じ向きであるためには関数$\boldsymbol{F}$のλにあたる変数の値が正であることが必要です。

したがって最適解の候補探索では次の連立不等式の解を求めることになります。

$$\mu>0,\ \ \nabla f(\boldsymbol{x})-\mu\nabla h(\boldsymbol{x})=\mathbf{0},\ h(\boldsymbol{x})=0,$$

$$\text{あるいは、}\nabla f(\boldsymbol{x})=\mathbf{0},\ h(\boldsymbol{x})>\mathbf{0}.$$

前者は境界上の探索、後者は内点の探索です。

複数の不等式制約のある最適化問題

次に、不等式制約が複数ある場合を考えます。

$$\min.\quad f(\boldsymbol{x}),\quad \text{s.t.}\qquad [h_1(\boldsymbol{x}),\ldots,h_l(\boldsymbol{x})]^{\mathrm{T}} \geqq \mathbf{0}.$$

複数の不等式制約を等号または不等号が成立するように満たす点の集合が実行可能領域になります。

単純に境界を定めることはできないので、まず不等式制約を等号で満たしているものと不等号で満たしているものに分類します。

具体的には、$l'(\leqq l)$個で等号が成立、残り$l-l'$個で不等号が成立したものとして、これを並べ変えたものを$h'_1,\ldots,h'_l$とし、これを有効制約関数と呼び、有効制約関数による制約を有効制約と呼びます。

$$\{h'_1,\ldots,h'_l\}=\{h_1,\ldots,h_l\},\ [h'_1,\ldots,h'_{l'}]^{\mathrm{T}}=\mathbf{0},\ [h'_{l'+1},\ldots,h'_l]^{\mathrm{T}}>\mathbf{0}.$$

複数の不等式制約のある最適化問題

境界上のゼロ点探索に対応する連立方程式は次の通りです。

$$\nabla f(\boldsymbol{x})-\Sigma_{i=1,\ldots,l'}\mu_i \nabla h'_i(\boldsymbol{x})=\mathbf{0},\ [h'_1(\boldsymbol{x}),\ldots,h'_{l'}(\boldsymbol{x})]^{\mathrm{T}}=\mathbf{0}.$$

未定係数にはさらに条件がつきます。

$\boldsymbol{x}$から内点方向に向かう微小ベクトル$\boldsymbol{d}$は次の条件を満たします。

$$\boldsymbol{d}^{\mathrm{T}}\nabla h'_i(\boldsymbol{x})\geqq 0 \quad (i=1,\ldots,l')$$

$\boldsymbol{d}$方向に向かうと目的関数が減少する条件

$$\boldsymbol{d}^{\mathrm{T}}\nabla f(\boldsymbol{x})<0$$

を加えて、両者を満たす$\boldsymbol{d}$が存在すると$\boldsymbol{x}$は目的関数が減少する内点領域に接していることになり、最適解の候補となりません。

複数の不等式制約のある最適化問題

前頁までの説明を整理し直すと、有効制約$[h'_1,\dots,h'_{l'}] \geqq \mathbf{0}$を等号で満たす境界上の点$\boldsymbol{x}$が最適解の候補であるためには、連立方程式

$$\nabla f(\boldsymbol{x}) - \sum_{i=1,\dots,l'} \mu_i \nabla h'_i(\boldsymbol{x}) = \mathbf{0},\ [h'_1(\boldsymbol{x}),\dots,h'_{l'}(\boldsymbol{x})]^{\mathrm{T}} = \mathbf{0}.$$

を$\boldsymbol{x}$が満たし、次の条件を満たす微小ベクトル$\boldsymbol{d}$が存在しないことが必要ということになります。

$$\boldsymbol{d}^{\mathrm{T}} \nabla h'_i(\boldsymbol{x}) \geqq 0 \quad (i=1,\dots,l'),\ \boldsymbol{d}^{\mathrm{T}} \nabla f(\boldsymbol{x}) < 0.$$

さらにこれらの条件を整理して得られる次頁の条件式を、その発見に関わった研究者の名前からKKT条件と呼びます。

※ KKT=Karush-Kuhn-Tucker

複数の不等式制約のある最適化問題

KKT条件（不等式制約のみを持つ場合）：

最小化問題 $\min. \quad f(\boldsymbol{x}), \quad \text{s.t.} \qquad [h_1(\boldsymbol{x}),\ldots,h_l(\boldsymbol{x})]^{\mathrm{T}} \geqq \mathbf{0}.$

において、次の条件（KKT条件）を満たす$\mu'_1,\ldots,\mu'_{l'}$が存在するとき
$\boldsymbol{x}$は有効制約$[h'_1(\boldsymbol{x}),\ldots,h'_{l'}(\boldsymbol{x})]^{\mathrm{T}}=\mathbf{0}$による境界上の最適解である。

$$\nabla f(\boldsymbol{x}) - \Sigma_{i=1,\ldots,l'} \mu_i \nabla h'_i(\boldsymbol{x}) = \mathbf{0},$$

$$\mu_i h'_i(\boldsymbol{x}) = 0,\ \mu_i \geqq 0,\ h'_i(\boldsymbol{x}) \geqq \mathbf{0} \quad \mathrm{i}=1,\ldots,\mathrm{l}'.$$

等式制約と不等式制約のある最適化問題

簡単のため、等式･不等式制約を1つずつ持つ場合を考えます。

最小化問題 $\min. \quad f(\boldsymbol{x}), \quad \text{s.t.} \qquad g(\boldsymbol{x})=0,\ h(\boldsymbol{x})\geqq 0.$

等式制約を合せて同等の制約となる2つの不等式制約に分割、

$$g(\boldsymbol{x})=0 \Leftrightarrow g^{+}(\boldsymbol{x})=g(\boldsymbol{x})\geqq 0,\ g^{-}(\mathrm{x})=-g(\boldsymbol{x})\geqq 0,$$

問題を3つの不等式制約を持つものに書き換えます。

最小化問題 $\min. \quad f(\boldsymbol{x}), \quad \text{s.t.} \qquad g^{+}(\boldsymbol{x})\geqq 0,\ g^{-}(\boldsymbol{x})\geqq 0,\ h(\boldsymbol{x})\geqq 0.$

不等式制約のみの問題なので、既に述べたKKT条件を書き下すことができます。（次頁）

※　$g^{+},\ g^{-}$は常に有効制約関数であることに注意します。

等式制約と不等式制約のある最適化問題

KKT条件：以下を満たす$[\lambda^+,\ \lambda^-,\ \mu]\geqq\mathbf{0}$が存在すれば、$\boldsymbol{x}$は境界上の最適解である。

$$\nabla f(\boldsymbol{x})-\lambda^+\nabla g^+(\boldsymbol{x})-\lambda^-\nabla g^-(\boldsymbol{x})-\mu\nabla h(\boldsymbol{x})=\mathbf{0},$$

不要 $\cancel{\lambda^+g^+(\boldsymbol{x})=0,\ \lambda^-g^-(\boldsymbol{x})=0},\ \mu h(\boldsymbol{x})=0,$

$$\lambda^+\geqq 0,\ \lambda^-\geqq 0,\ \mu\geqq 0,$$

$g(\boldsymbol{x})=0$
に置き換える $\cancel{g^+(\boldsymbol{x})\geqq 0,\ g^-(\boldsymbol{x})\geqq 0},\ h(\boldsymbol{x})\geqq 0.$

$g^+,\ g^-$の定義と最後の不等式から$g^+(\boldsymbol{x})=g^-(\boldsymbol{x})=g(\boldsymbol{x})=0$なので2行めの等式は$\lambda^+,\ \lambda^-$に依らないことが判ります。

等式制約と不等式制約のある最適化問題

（前頁の続き）

KKT条件：以下を満たす$[\lambda^{+},\ \lambda^{-},\ \mu]\geqq\mathbf{0}$が存在すれば、$\boldsymbol{x}$は境界上の最適解である。

$(\lambda^{+}-\lambda^{-})\nabla g(\boldsymbol{x})$に置き換える

$$\nabla f(\boldsymbol{x})\cancel{-\lambda^{+}\nabla g(\boldsymbol{x})-\lambda^{-}\nabla g(\boldsymbol{x})}-\mu\nabla h(\boldsymbol{x})=\mathbf{0},$$

$$\mu h(\boldsymbol{x})=0,$$

$$\lambda^{+}\geqq 0,\ \lambda^{-}\geqq 0,\ \mu\geqq 0,$$

$$g(\boldsymbol{x})=0,\ h(\boldsymbol{x})\geqq 0.$$

第1式を定義より$\nabla f(\boldsymbol{x})-(\lambda^{+}-\lambda^{-})\nabla g(\boldsymbol{x})-\mu\nabla h(\boldsymbol{x})=\mathbf{0}$ と書き換えると、係数をまとめて扱うことができます。

等式制約と不等式制約のある最適化問題

（前頁の続き）

KKT条件：以下を満たす$[\lambda^+, \lambda^-, \mu] \geqq \mathbf{0}$が存在すれば、$\boldsymbol{x}$は境界上の最適解である。

$$\nabla f(\boldsymbol{x}) - (\lambda^+ - \lambda^-)\nabla g(\boldsymbol{x}) - \mu \nabla h(\boldsymbol{x}) = \mathbf{0},$$

$$\mu h(\boldsymbol{x}) = 0,$$

不要 $\cancel{\lambda^+ \geqq 0, \lambda^- \geqq 0,}\ \mu \geqq 0,$

$$g(\boldsymbol{x}) = 0,\ h(\boldsymbol{x}) \geqq 0.$$

係数$(\lambda^+ - \lambda^-)$を正とするときは$\lambda^- = 0$、負とするときは$\lambda^+ = 0$でλ^+, λ^-の非負条件を満たすことができるので非負条件は無意味です。

等式制約と不等式制約のある最適化問題

（前頁の続き）

KKT条件：以下を満たす$[\lambda^+,\ \lambda^-,\ \mu]\geqq\mathbf{0}$が存在すれば、$\boldsymbol{x}$は境界上の最適解である。

λと書き換えてよい

$$\nabla f(\boldsymbol{x})-\cancel{(\lambda^+-\lambda^-)}\nabla g(\boldsymbol{x})-\mu\nabla h(\boldsymbol{x})=\mathbf{0},$$

$$\mu h(\boldsymbol{x})=0,\ \mu\geqq0,\ g(\boldsymbol{x})=0,\ h(\boldsymbol{x})\geqq0.$$

上記の通りKKT条件が整理された。係数$(\lambda^+-\lambda^-)$は1つの係数なので、これを記号λで置き換えることができる。

この結果を一般化すると複数の等式・不等式制約のある最適化問題におけるKKT条件を得ることができる。（次頁）

等式制約と不等式制約のある最適化問題

（前頁の続き）

KKT条件：以下を満たす$[\mu_1,\ldots,\mu_l]\geqq\mathbf{0}$が存在すれば、$\boldsymbol{x}$は境界上の最適解である。

$$\nabla f(\boldsymbol{x})-\Sigma_j\lambda_j\nabla g_j(\boldsymbol{x})-\Sigma_k\mu_k\nabla h_k(\boldsymbol{x})=\mathbf{0},$$

$$g_j(\boldsymbol{x})=0,\ h_k(\boldsymbol{x})\geqq0,\ \mu_k\geqq0,\ \mu_k h_k(\boldsymbol{x})=0,$$

$$j=1,\ldots,m,\ k=1,\ldots,l.$$

一般の制約のある最適化問題の最適性条件が得られた。

最適解の探索

一般の制約付き最適化問題の最適性条件が得られたので、最適解を実際に得るためには最適性条件を満たす変数値の組み合わせを探索する必要があります。

すでに説明された降下法やニュートン法に属する方法が利用できますが、一般に解が得られることを保証することはできません。

これは、初期解の設定等の適用の仕方に因る事情もあり得ますが、そもそも与えられた問題から自然に得られる最適性条件を満たす解の安定性のような、適用法によって解決できない問題も考えられます。

そこで、例えば収束性や解の安定性を向上するために問題あるいは最適性条件を構成する関数の置き換えをする方法があります。

最適解の探索

罰則（ペナルティ）法

線形計画問題で扱ったように制約（線形計画問題では人工問題において人工変数値がゼロであるという等式制約でした）の非実現に大きな重みを置く方法です。

例えば、最小化問題の等式制約　$g(\boldsymbol{x})=0$ に対して大きな定数Pをかけて目的関数に加えるというのが線形計画問題で紹介した罰則法でした。

$$\min.\ f(\boldsymbol{x}) \Rightarrow f(\boldsymbol{x})+Pg(\boldsymbol{x})$$

変更後の問題は元の問題と同じ最適解を持ち、また制約を満たす最適解において最適性条件を満たします。

最適解の探索

（前頁の続き）罰則（ペナルティ）法

一方で、適切な罰則定数を事前に決められない、Hesse行列の正定値性に影響があること等が問題となります。

そこで、罰則の付与を目的関数ではなく最適性条件に与えるという方法が使われています。

KKT条件に現れるベクトル関数FはLagrange関数と呼ばれる目的関数に制約関数を組み合わせた関数の勾配ベクトルと言えます。（行列・ベクトル表現に注意してください。）

Lagrange関数：　$L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})=f(\boldsymbol{x})-\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x})$

最適解の探索

（前頁の続き）罰則（ペナルティ）法

Lagrange関数に現れる制約関数に対して罰則関数を加え、結果として反復計算で扱う最適性条件のゼロ点探索において罰則の効果を得るというものです。

この方法は乗数法と呼ばれています。

最適解の探索

障壁（バリア）関数

非負条件を課された制約関数を置き換える関数で、罰則（ペナルティ）と似た考え方で境界で極端な値が出るものを使います。

例えば不等式制約　$h(\boldsymbol{x}) \geqq 0$　の制約関数を

$1/h(\boldsymbol{x})$や$-\log h(\boldsymbol{x})$

のような境界で発散するものに置き換えて最小化問題の目的関数に加えたりLagrange関数を置き換えるという使い方をします。

また、線形計画問題の内点法において最小化の過程でたどる経路（パス）が実行可能領域の内点にあることを保証するために使われます。

まとめ

第一部：線形計画問題の最適化法

第二部：制約の無い非線形計画問題の最適化法

第三部：制約のある非線形計画問題の最適化法

という三部構成で授業（解説）をしました。

多くの分野で必要な連続関数の最適化について、基本的なことを簡潔に説明するという趣旨で用意した内容です。

実用的な計算法という意味では、各部とも十分とは言えず追加の勉強が必要です。

一方で、有用なツールと言える計算ライブラリが様々な形で利用できますので、授業で得た知識をもとに活用するようなことを考えてください。

まとめ

（前頁の続き）

最適化法について興味を持ち、実用的な計算法、あるいは最近の研究状況について知りたいと思われた方は、協力したいと思いますので、是非担当教員（2020年度は岡野）まで申し出てください。

演習問題1

次の3つの関数を目的関数とした制約なし最小化問題を考える。

(A) $x_1{}^2-2x_1x_2$, (B) $x_1{}^2-x_2{}^2$, (C) $2x_1{}^2+x_2{}^2$,

問1:各目的関数の最急降下ベクトルを示せ。

問2:初期点 $x_1=1,\ x_2=1$から最急降下法を用いて解くことを考える。最初の 1 ステップを解き、更新後の変数値を示せ。直線探索では厳密な解が解けたものとして良い。

(ヒント)

初期点を$\boldsymbol{x}_0$とすれば最急降下法の次のステップの近似解は$\boldsymbol{x}_1=\boldsymbol{x}_0-t\nabla f\,(\boldsymbol{x}_0)$と表すことができます。直線探索問題は $f(\boldsymbol{x}_1)=f(\boldsymbol{x}_0-t\nabla f\,(\boldsymbol{x}_0))$の最小化問題です。

この問題で示された(A)~(C)であれば、1変数tの停留点を求めることは易しいので近似計算をする必要はありません。

演習問題2

次の3つの関数を目的関数とした制約なし最小化問題を考える。

(A) ${x_1}^2 - 2x_1x_2$,　(B) ${x_1}^2 - {x_2}^2$,　(C) $2{x_1}^2 + {x_2}^2$,

問1：各目的関数のニュートン方向ベクトルを示せ。

問2：初期点 $x_1 = 1,\ x_2 = 1$から最急降下法を用いて解くことを考える。最初の 1 ステップを解き、更新後の変数値を示せ。

（演習問題1をニュートン法に置き換えたものです。）

演習問題3

次の最小化問題にLagrangeの未定係数法を適用し計算過程とともに最小解を示せ。

$$\min.\quad f(x_1,\ x_2)=ax_1+bx_2+c,\ \text{s.t.}\ \ g(x_1,\ x_2)=x_1{}^2+x_2{}^2=1.$$

ヒントと注意：

例題の制約式と目的関数を入れ替えた問題です。例題の図から最適解は明らかでしょう。

停留点が複数あることに注意してください。

演習問題4

次の最小化問題のKKT条件を示し、最適解を求めよ。

$$\min. \quad f(x_1,\ x_2) = {x_1}^2 + {x_2}^2,\ \text{s.t.}\ h(x_1,\ x_2) \geqq ax_1 + bx_2 + c.$$

ヒントと注意:

これも例題の変形です。

a, b, c の値による影響も考慮してください。

煩雑であれば、$a = b = c = 1$ の場合だけを考えても良いです。