2020年度「数値最適化」

第二部「制約の無い一般の最適化問題」
(1)変数が1つの場合

非線形計画問題

授業の前半では数理計画問題を構成する目的関数や制約式を1次の線形多項式に限定し、そのことを利用した方法について学習しました。

一般にそのような問題を線形計画問題と呼びます。

授業の後半は前半で設けたような限定を取り除き一般の数値最適化について扱います。

こうした一般の数理計画問題を、線形計画問題と区別して、非線形計画問題と呼び、その方法を非線形計画法と呼ぶことがあります。

これは「非線形の場合だけを扱う」のではなく「線形である」という条件なしで問題を解こうということです。

非線形計画問題の定式化

授業の後半で扱う「非線形計画問題」の定式化について述べる前に、一般の数理計画問題を次のように表現します。

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

すなわち、n変数の目的関数fを最小化するような変数の値$x_1,\ldots,x_n$をm個の等式、l個の不等式を満たすもののなかから見つける問題を考えます。

このとき制約を与える$g_1,\ldots,g_m,h_1,\ldots,h_l$を制約関数と呼びます。

非線形計画問題の定式化

授業の前半で用いたベクトル等式不等式の表現を流用すれば、

$$\min.\quad f(\boldsymbol{x}),\ \text{s.t.}\ \boldsymbol{G}(\boldsymbol{x}){=}\boldsymbol{0},\ \boldsymbol{H}(\boldsymbol{x}){\geqq}\boldsymbol{0}. \tag{1}$$

と書くことができます。

ただし、$\boldsymbol{x}{=}[x_1,\ldots,x_n]^{\mathrm{T}}, \boldsymbol{G},\ \boldsymbol{H},\ \boldsymbol{0}$は長さ$n$のベクトルおよびベクトル関数、ゼロベクトルで等号･不等号は左右両辺のベクトルの全ての成分が対応する対において等式･不等式を満たすことを表します。

関数の最大化や向き等の異なる不等号により、(1)式で直接に表すことのできない問題を想定することも可能ですが「同等の問題」を考えることで(1)の形式の問題を扱うことで全般的な問題を扱えることは線型計画問題の標準形において十分に説明しました。

非線形計画問題の定式化

数理計画問題　$\min. \quad f(\boldsymbol{x}), \text{ s.t. } \boldsymbol{G}(\boldsymbol{x})=\boldsymbol{0},\ \boldsymbol{H}(\boldsymbol{x})\geqq\boldsymbol{0}. \qquad (1)$

において関数$f, \boldsymbol{G},\ \boldsymbol{H}$を$\boldsymbol{x}$の線形関数に限定したものが「線形計画問題」と言えます。

関数$f, \boldsymbol{G},\ \boldsymbol{H}$を$\boldsymbol{x}$の線形関数に限定しない問題を「一般の最適化問題」あるいは「非線形計画問題」と呼びますが、授業の後半では

- 関数$f, \boldsymbol{G},\ \boldsymbol{H}$は$\boldsymbol{x}$の連続関数である
- 関数$f, \boldsymbol{G},\ \boldsymbol{H}$の1階偏導関数を定められ、それは連続である
- 関数$f, \boldsymbol{G},\ \boldsymbol{H}$の2階偏導関数を定められ、それは連続である

という仮定を必要に応じて採用し、「非線形計画問題」とします。

制約の無い非線形計画問題（1変数の場合）

制約の無い非線形計画問題：　$$\min.\quad f(\boldsymbol{x}). \tag{2}$$

制約関数の無い自由変数$\boldsymbol{x}$の非線形計画問題（2）を考えます。

例えば、次の多項式

$$f(x) = 9x^4 - 2x^3 - 18x^2 + 6x$$

の最小値を与える変数xの値を求める問題はベクトル$\boldsymbol{x}$の長さが1、あるいは変数の個数が1の「制約の無い非線形計画問題」と言えます。

中学校あるいは高等学校の数学の授業では、このような問題をどのように扱ったかを思い出してください。

制約の無い非線形計画問題（1変数の場合）

「数学II」では、まず多項式$f(x)$から導関数$f'(x)$を求めます。

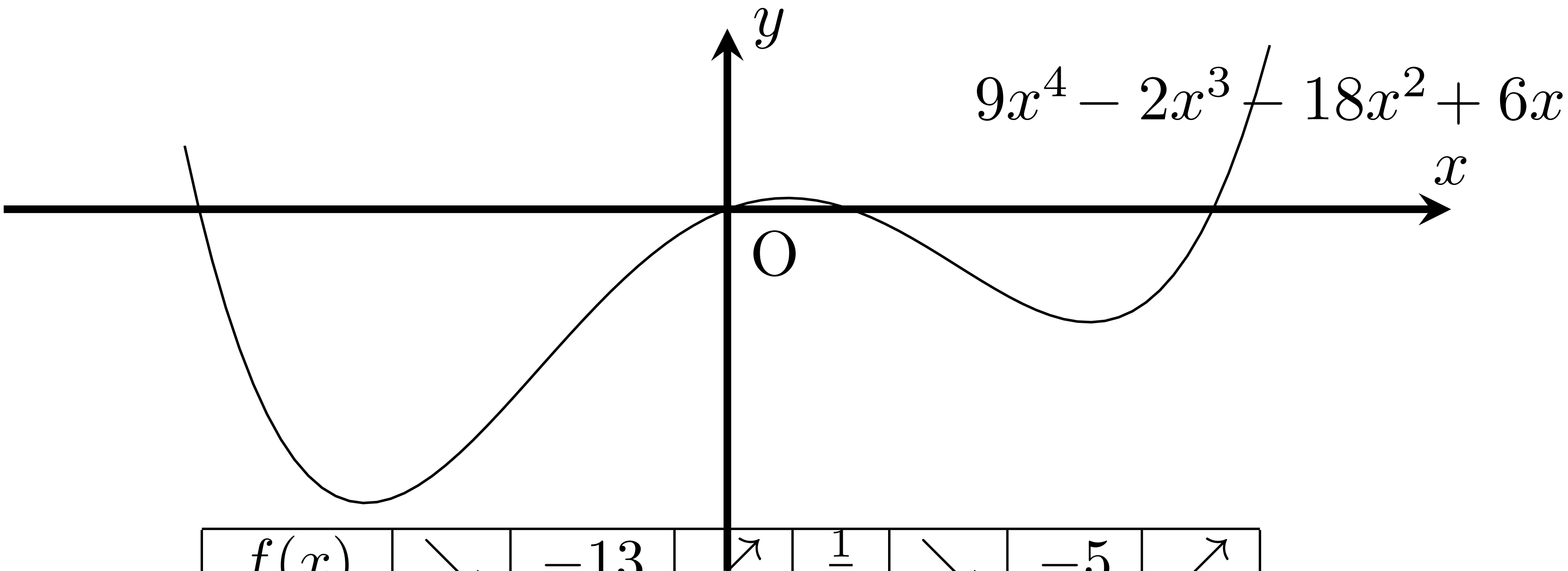
$$f'(x)=36x^3-6x^2-36x+6$$

$f'(x)=0$を満たす全てのゼロ点を求め、ゼロ点で定義域を区切れば、区間毎に導関数の符号を定め、$f(x)$の増減表が得られます。

増減表を見れば最小値の候補が見つかり、その中から最小値と、最小値を与えるxを決めることができます。

$f(x)$	$\searrow$	-13	$\nearrow$	$\frac{1}{6}$	$\searrow$	-5	$\nearrow$
$f'(x)$	$-$	0	$+$	0	$-$	0	$+$
x		-1		$\frac{1}{6}$		$+1$	

制約の無い非線形計画問題（1変数の場合）



$f(x)$	$\searrow$	-13	$\nearrow$	$\frac{1}{6}$	$\searrow$	-5	$\nearrow$
$f'(x)$	$-$	0	$+$	0	$-$	0	$+$
x		-1		$\frac{1}{6}$		$+1$	

制約の無い非線形計画問題（1変数の場合）

「数学II」で習った方法は$f(x)$の導関数$f'(x)$を決めることができて、fとf'が連続であることが前提になっています。これは授業の後半における条件とも一致します。

x^*が$f(x)$の最小解で$f(x^*)$が最小値であるなら、任意のdで

$$f(x^* + d) \geqq f(x^*)$$

が成立するはずです。

このとき、$f'(x^*)$を1つの値に決めるためには$f'(x^*) = 0$が必要です。

したがって、$f'(x^*) = 0$はx^*が最小解であるための必要条件です。

制約の無い非線形計画問題（1変数の場合）

関数値が0をとる点の変数値をその関数のゼロ点と呼びます。

最小解x^*は$f'(x^*)=0$を満たすのでf'のゼロ点です。

導関数のゼロ点を元の関数の停留点と呼びます。

最小解x^*は$f'(x^*)=0$を満たすのでfの停留点です。

f'のゼロ点であることとfの停留点であることは同じです。

x^*がfの停留点であること（もしくはf'のゼロ点であること）はx^*が最小解であるための必要条件ですから、最適解の探索手順として、まず停留点を探索することが考えられます。

制約の無い非線形計画問題（1変数の場合）

停留点であるだけでは最適解とは言えませんが最適解は停留点に含まれているので、全ての停留点を求め、その関数値を比較すれば最適解を定めることができます。

したがって、$f'(x^*)=0$はx^*が最小解であるための必要条件です。

少し前のスライドに示した多項式の例では導関数の因数分解により、全ての停留点を求めることができました。

線形計画問題の最初に説明した総当たりの解法に似ているかもしれません。

制約の無い非線形計画問題（1変数の場合）

一般の非線形計画問題において、f'の全ての停留点を求める方法は非効率で現実的ではありません。

線形計画問題で総当たりの方法が変数や制約条件の多い実際的な問題に不適だったのと同様です。

そこでまずは最小解の候補となる停留点を1つ見つける方法を考えます。

これは方程式$f'(x^*)=0$の解x^*を探索することなので、種々の逐次反復法が利用できます。

すなわちx^*に近い初期値x_0を順次改善して得られる数列$x_0, x_1, \ldots$から方程式を満たす、あるいは近似するx_Nを見出す方法です。

制約の無い非線形計画問題（1変数の場合）

方程式の逐次解法には様々なものがあります。ここでは、関数$f(x)$の最小化という目的に即して次の更新式を採用します。

$$x_{k+1} = x_k - r_k f'(x_k) \ (r > 0)$$

変数の更新量$|r_k f'(x_k)|$を微少量と見做せる範囲であれば

$$f(x_{k+1}) \approx f(x_k) - r_k [f'(x_k)]^2 < f(x_k)$$

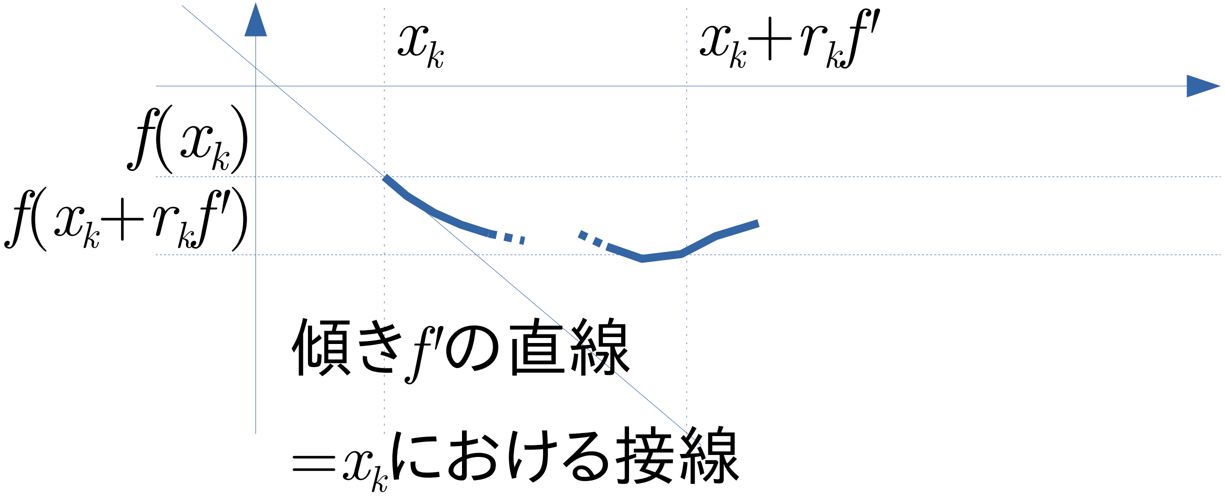
となり、目的関数の減少を期待できます。

これをくり返し、どこかで$f'(x_N) = 0$あるいはそれに近い条件$|f'(x_N)| << 1$を実現したところで停止することができます。

制約の無い非線形計画問題（Armijoの方法）

前ページで採用した更新式にはr_kという未定係数があります。実際に計算を進めるためには、これを決める必要があります。

r_kを大きくとれば目的関数の減少は大きくなりそうですが、停留点や最小点を通り過ぎてしまうかもしれません。

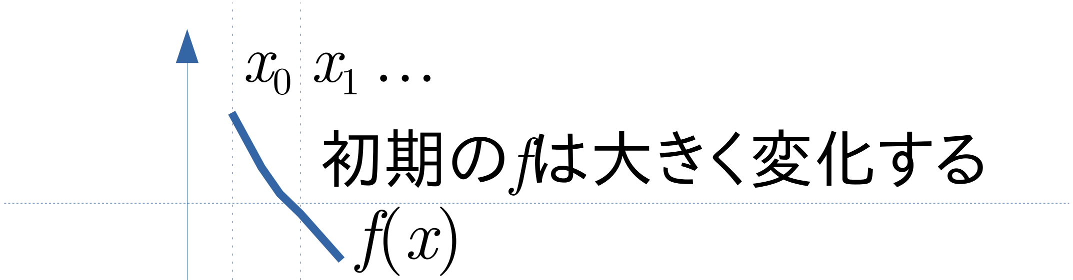


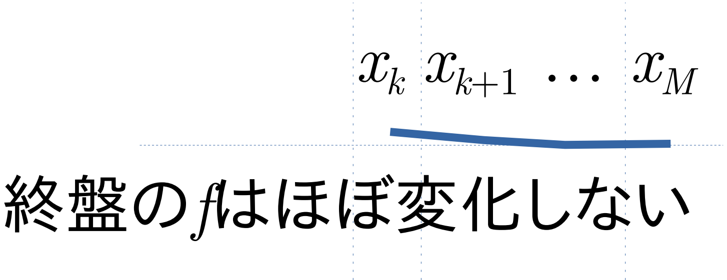
制約の無い非線形計画問題（Armijoの方法）

微分可能な連続関数fが初期点x_0近辺に停留点$f'(x^*)=0$を持つという前提なので、点列$x_0, x_1, \ldots,$を細かくとればx^*の良い近似に当たるはずです。

とはいえ、更新量$x_{k+1}-x_k$が小さければ探索の完了までの計算回数は膨大になります。早くx^*を得るために点列は粗く採るべきです。

また、f'が0に近づく点を探索しているので、fの傾きも大きく変化しますから、点列の間隔を固定するのは難しくなります。

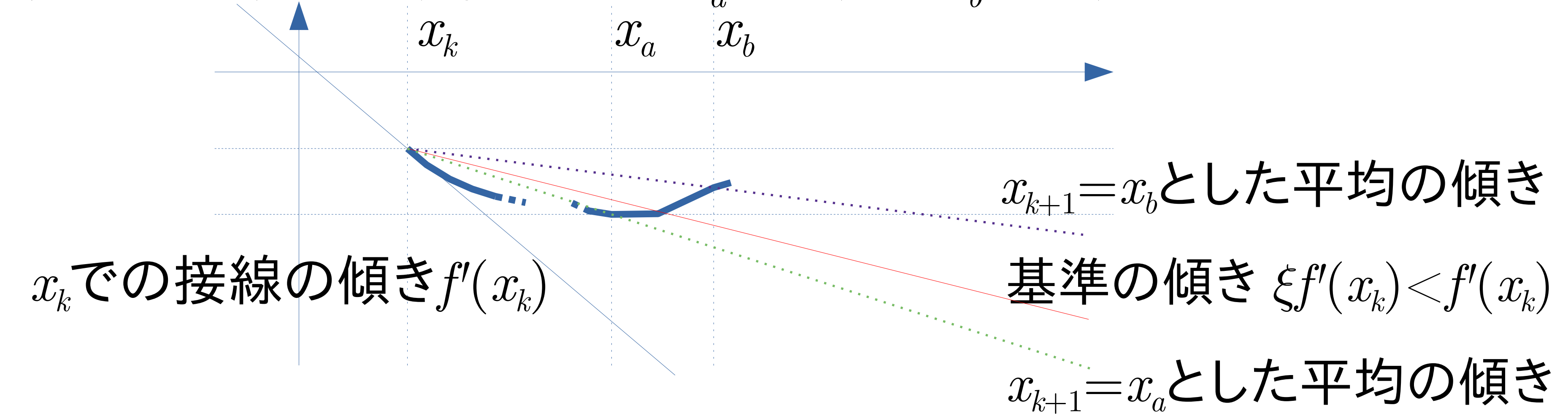




制約の無い非線形計画問題（Armijoの方法）

そこでr_kや変数の更新量$x_{k+1}-x_k$を固定するのではなく更新毎の（平均の）傾きを基準に更新量を決める方法を採用します。

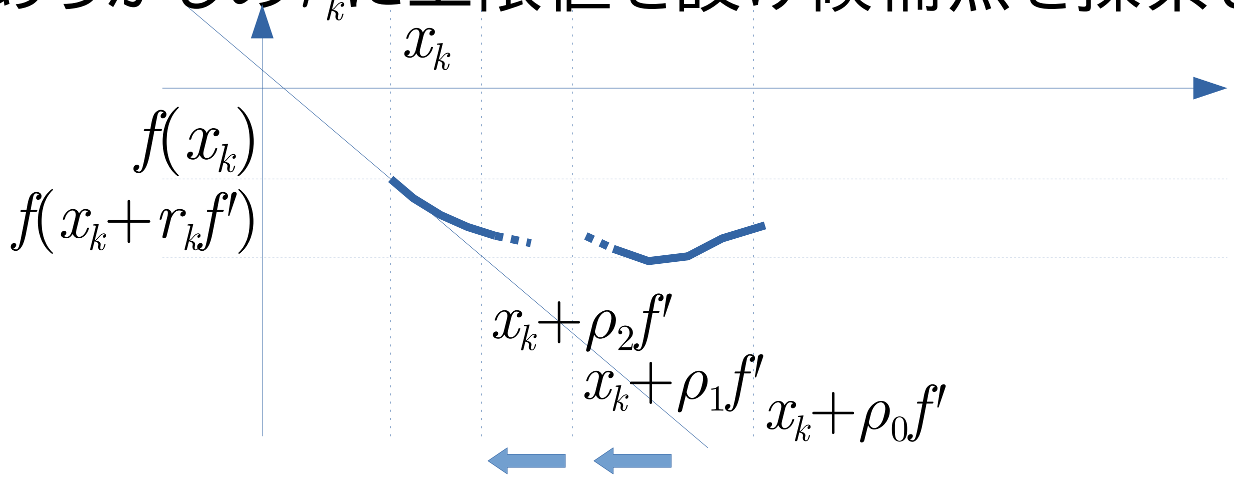
具体的には定数$0<\xi<1$を定め$x_k \rightarrow x_{k+1}$の平均の傾きとx_kにおける接線の傾きとの比がこの定数を下回らないようにx_{k+1}を定めるという方法を採用します。（∴図のx_aを採用しx_bを採用しない）



制約の無い非線形計画問題（Armijoの方法）

前ページで採用した平均の傾きによる更新をアルミホ（Armijo）の方法と呼び更新量の採用条件をアルミホ条件あるいは規準と呼びます。

アルミホ条件のみで更新を進めるには2分法を採用するのが現実的でう。あらかじめr_kに上限値を設け候補点を探索します。（図）

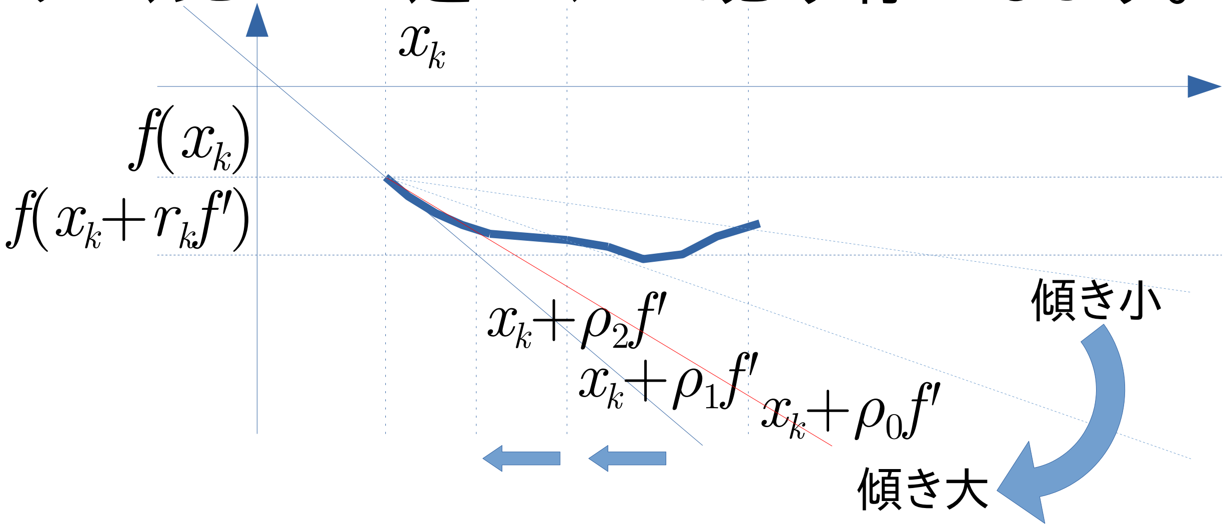


制約の無い非線形計画問題（Armijoの方法）

下図のようにρ_0から順にρ_kの候補を求め、探索を進めます。

x_{k+1}の候補点における平均の傾きとx_kにおける接線の傾きf'との比に1以下の上限を定めて、2分法の停止条件とすることができます。

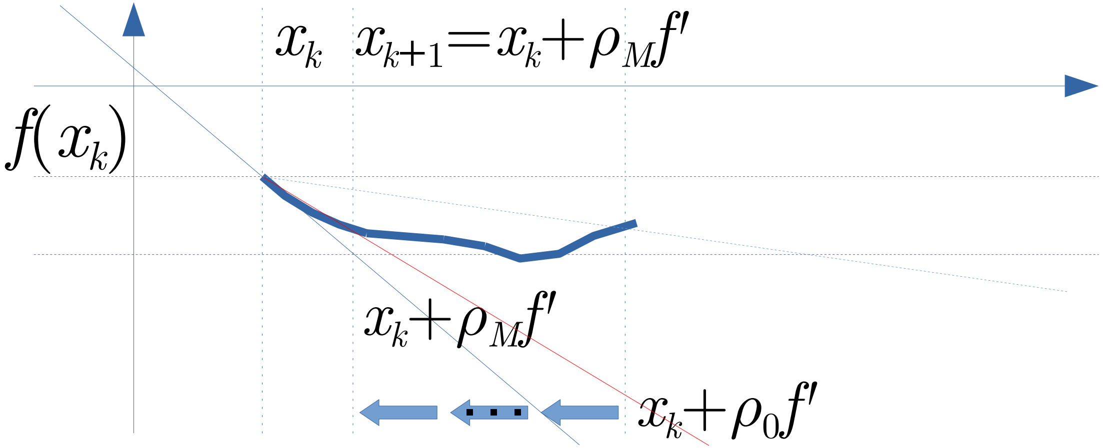
x_kに近づけば、比は1に近づくので必ず停止します。



制約の無い非線形計画問題（Armijoの方法）

典型的な2分法（$\rho_k/\rho_{k+1}=2$）でなくても構いませんが、いずれにしても1回の更新でおおよそξの公比で等比級数的に傾きを減少させて停留点を見つけようとする方法と言えます。

結果の良否はパラメタξ次第ということもできます。



制約の無い非線形計画問題（Wolfeの条件）

前ページまでで説明した方法は変数xの代わりに傾きを使って1次元の探索をするものです。

実際には$r_k{=}\rho_0,\ \xi$のどちらかの比で傾きが更新されますが、これは平均の傾きに注目した更新なので停留条件$f'{=}0$を基準にするものではありません。

そこで、更新後の導関数値による次のような規準も考えられます。

$$\omega f'(x_k) < f'(x_k + r_k f'(x_k)) \quad \omega < 1$$

すなわち、導関数値自体が公比ωの等比級数より速く0に近づくことを要請するものです。これをWolfeの条件と呼びます。

降下法

Armijo条件やWolfe条件以外にも様々な規準（条件）を想定することができます。また、具体的な計算手順も2分法だけでなく、多くの異なる方法を考えることができるでしょう。

前ページまでに説明した方法の基本的なアイディアは、

- $f'(x_k)$の符号によって更新の向き（符号）を決める
- $f(x_{k+1})<f(x_k)$となり目的関数が減少するように更新する

というものです。このような方法を総称して降下法と呼びます。

停留点と最小解・極小解

微分可能な目的関数の最適解が停留点であることは必要条件ではありますが、十分条件ではありません。

すなわち、降下法等を用いて得た停留点は最小解の候補ではありますが、必ずしも最小解ではありません。

数学IIでは増減表あるいは導関数の符号を用いて停留点前後での関数値の増減を求め極小解であることを判定しました。

このことを、もう少し厳密に考え直し、また一般性のある形式で定義し直して多変数関数の問題にも利用できるよう準備します。

停留点と最小解･極小解

凸関数

次の不等式を満たす関数を凸関数と呼びます。

任意の$x,y,0 \leqq t \leqq 1$ に対し　$tf(x)+(1-t)f(y) \geqq f(tx+(1-t)y)$

※定義域が限定される場合はx,yは定義域からとります

狭義凸関数

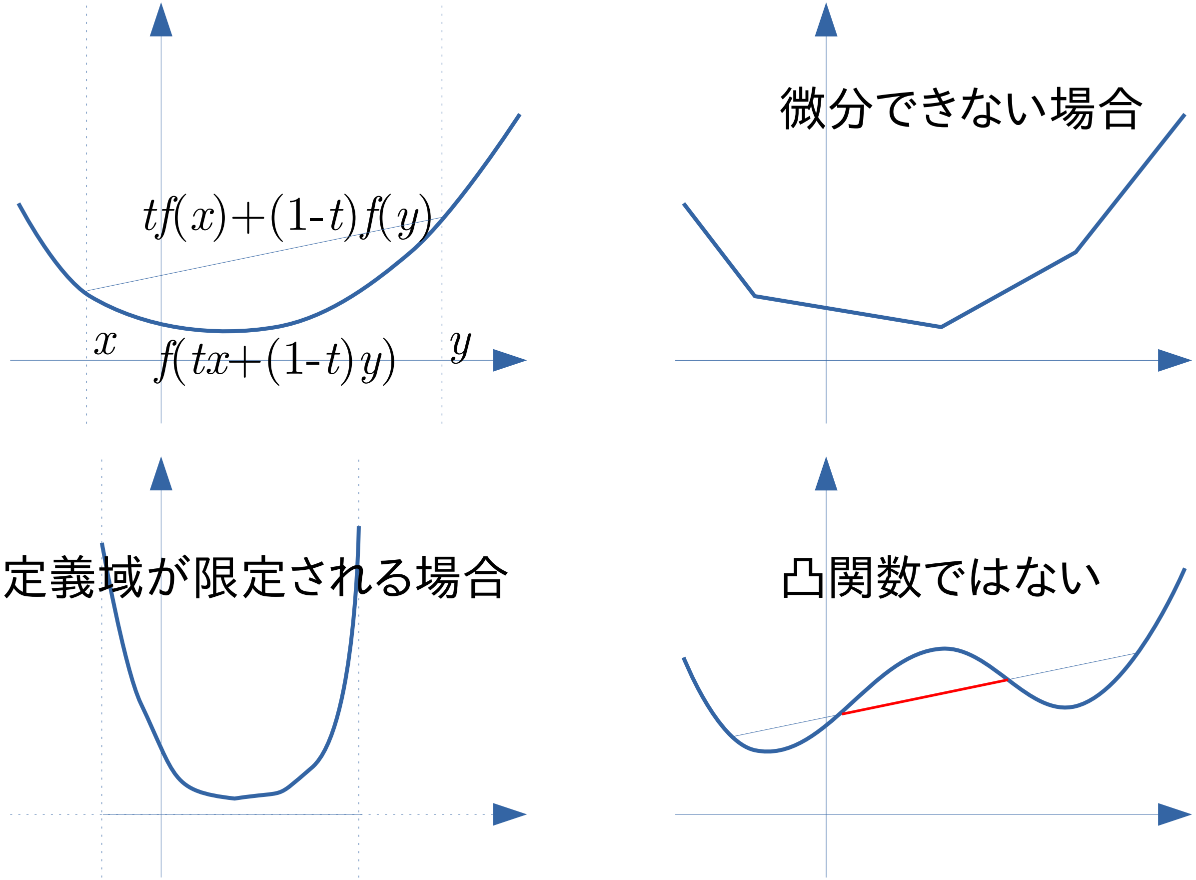
次の不等式を満たす関数を狭義凸関数と呼びます。

任意の$x,y,0 \leqq t \leqq 1$ に対し$tf(x)+(1-t)f(y) > f(tx+(1-t)y)$

※条件式の不等号に等号が含まれません

停留点と最小解・極小解

凸関数の例

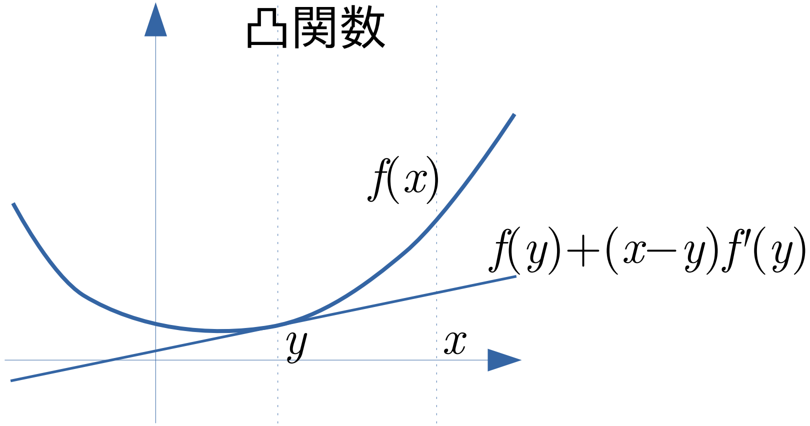


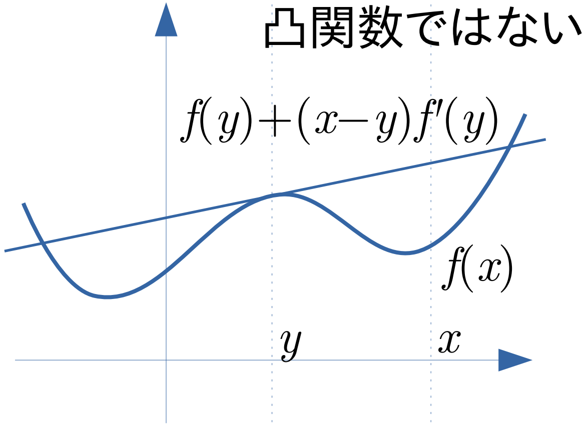
停留点と最小解・極小解

1階微分可能な凸関数の定理

次の不等式は1階連続微分可能な凸関数の必要十分条件です。

任意のx,yに対し　$f(x) \geqq f(y) + (x-y)f'(y)$





停留点と最小解･極小解

凸関数の停留点について大域的最適性が成立します。

定理：凸関数fの停留点は大域的最適解（＝最小解）である

fは凸関数なので、任意のx,yに対し

$$f(x) \geqq f(y) + (x - y) f'(y)$$

x^*が停留点であれば$f'(x^*) = 0$

したがって、任意のxについて

$$f(x) \geqq f(x^*) + (x - x^*) f'(x^*) = f(x^*)$$

すなわちx^*は最小解となる。

停留点と最小解･極小解

凸関数の停留点の局所的最適性（極小解）の定理

fが2階微分可能であれば2次のTaylor展開を考えられる

$f(x+d)=f(x)+df'(x)+(d^2/2)f''(x)+\mathrm{o}(|d|^2)$

x^*が停留点であれば$f'(x^*)=0$、したがって十分に小さなdについて

$$f(x)≧f(x^*)+(d^2/2)f''(x^*)$$

とすることができるので$f''(x^*)>0$ならばx^*はその近傍において極小（極小解）と言える。

制約の無い1変数の非線形計画法

まとめ

1変数の目的関数による（制約の無い）非線形計画問題では、

数学IIで学ぶ増減表を用いた最適解探索の方法から出発して、一般の非線形計画問題に適用できる降下法による最適解の探索法を学びました。

1変数の降下法は多変数に拡張することができるだけでなく、多変数の降下法中で重要な役割りを果たす直線探索の方法となります。

目的関数の凸性を定義し、この性質が停留点の大域最適性、あるいは局所最適性に関わることを理解しました。