経済数学II（秋学期）

服部哲弥

START ↱

秋学期初回講義（教科書p.27）の前に

○　クラス指定なので秋のガイダンスは「春学期（経済数学１）の続き」だけですが，念のため春学期初回の中から

- 標準のテキスト（水色）「経済数学」の続きで秋学期はII５－７章
- テキストの春学期の内容（および１年の数学の内容）は既知とする

○　学事日程上毎回（ほぼ毎週）講義曜日に講義pdfをDL（今回が1）

- 毎回の講義pdfに問のページ，次の回の講義pdfに略解のページ

○　**成績評価**

- 学期末定期試験が教室で行われれば**持ち込み無しの学期末定期試験**
- 万が一教室試験が有効に機能せずオンライン試験になれば，**教科書と上記の講義pdfを手持ちであることを前提に出題**
- 「社会情勢」の急変に応じて大学の方針の急変がありうるので，学事指定講義用オンラインサイトの本科目のお知らせの指示に注意し協力すること．試験等終了までは**ダウンロードした毎回の講義と教科書所持を続けること**

秋の内容

(1-c) **不等式条件付き最小値問題：**

（(1-a) **条件無し，**(1-b) **等式条件，は春に学習済み）**

$xy + x + y \leqq 1,\ x \geqq 0,\ y \geqq 0$ **の下で** xy **の最大値はやや難．**

２**変数なら領域の図示で頑張れるが，**３**変数以上だと？**

- Fritz–John **条件と** Kuhn–Tucker **条件＝未定乗数法と似た工夫の成果：**
- （２変数以下や特別な問題に限らない） **未定乗数法のような汎用方法で**
- **未定乗数法に**（したがって条件無しの時の $\nabla f(\overrightarrow{x}) = \overrightarrow{0}$ に）**似た公式**
- **そんな戦闘力高い公式を得る理論系経済学向きの汎用・強力な理屈が**

(2)　**凸集合の分離定理：**

- **「**２**個の膨らんだ風船の間に紙を挟める」という定理（汎用）**
- **特に，風船のいっぽうが線形空間の場合（**Gordan **の定理）（強力）**

○早速ですが，次のページから今回の講義なので，必ず次のページへ．

１．Fritz–John条件（教科書p.27）

m と n は正整数．n 変数 C^1 級実数値関数たち f, g_1, g_2, ..., g_m．
点 $\overrightarrow{a} \in \mathbb{R}^n$ は不等式条件 $g_1(\overrightarrow{a}) \leqq 0, \ldots, g_m(\overrightarrow{a}) \leqq 0$ を満たす．
以上を１組与えられたとして講義１，２，３で固定する．

Fritz–John条件：

$\exists \mu_0, \mu_1, \ldots, \mu_m;$ （←以下を満たす μ_i たちの組がある，の意味：春学期参照）

$$\mu_0 \nabla f(\overrightarrow{a}) + \mu_1 \nabla g_1(\overrightarrow{a}) + \cdots + \mu_m \nabla g_m(\overrightarrow{a}) = \overrightarrow{0},$$

$$\mu_1 g_1(\overrightarrow{a}) = \cdots = \mu_m g_m(\overrightarrow{a}) = 0,$$

$$\mu_0 \geqq 0,\ \mu_1 \geqq 0,\ \ldots,\ \mu_m \geqq 0,$$

$$(\mu_0, \mu_1, \cdots, \mu_m) \neq \overrightarrow{0}.$$

∇f は f の勾配ベクトル（春学期の講義参照）

不等式条件下の極小値（教科書p.27第5章定理2）

定理（Fritz–John条件）．前ページで固定した関数たち $f, g_1, \ldots, g_m$ について，**条件 $\overrightarrow{g} \leqq 0$ の下での極小値を f が $\overrightarrow{a}$ で取るならば**，前ページの $\overrightarrow{a}$ でのFritz-John条件（以下講義内略記で**FJ条件）が成り立つ**．

注．不等式条件下なので，春学期の等式条件下の極値問題と違って，たとえば，f の極小値と f の極大値は条件の符号や不等号が部分的に変わる．よって，

- f_1 の**極大値問題は $f = -f_1$ に対する極小値問題に直して**から教科書のFJ条件を使い，**解（の候補）を得てから符号を逆にして**元の f_1 の極大値（最大値）として答える．また，
- 条件がたとえば $g_1 \geqq g_2$ **の形のときは $g = g_2 - g_1$ として条件を $g \leqq 0$ に直して**教科書のFJ条件を使う．

極小値の必要条件の定理としての先人の工夫

勾配ベクトルについての方程式の比較（他の条件や場合分けの詳細は割愛）

条件無し： $\nabla f(\overrightarrow{a}) = \overrightarrow{0}$

等式条件下：$\nabla(f(\overrightarrow{a}) + \lambda g(\overrightarrow{a})) = \overrightarrow{0}$ 　　（未定乗数法）

不等式条件下：$\nabla(\mu_0 f(\overrightarrow{a}) + \mu_1 g(\overrightarrow{a})) = \overrightarrow{0}$ 　（ＦＪ条件）

（条件付は条件 $m = 1$ の例．微分の線形性を用い線形結合後に勾配ベクトルを取る形に等式変形済.）

- 昔の偉い人が工夫して上記のように**条件無しと似た形に整形**．
  - 等式条件の場合は等式で変数を消去して条件無しにしたとき，変数消去の代入で複雑になったのを整形（春学期講義ｐｄｆ１１参照）
  - 不等式条件 $g \leqq 0$ の下での f の極小の場合は，$g = 0$ の場合と $g < 0$ の場合に分けて考えると，それぞれ等式条件と条件無しになるが，場合分けで煩雑になるのをまとめたのがＦＪ条件（次ページに続く）

Gordanの定理からＦＪ条件へ

（前ページからの続き）$g=0$の場合と$g<0$の場合に分けて考えると，$g=0$は等式条件なので未定乗数法が利用可能（春学期既習）．
$g(\overrightarrow{a})<0$のとき？冒頭設定によって，gが（関係するすべての関数が）C^1級（以下春学期講義ｐｄｆ６参照）の場合なので，連続だから，$g(\overrightarrow{a})<0$ならば$\overrightarrow{a}$のある近傍でべたっと$g<0$が成立．つまり，$g\leqq 0$は変数の取り得る値を$\overrightarrow{a}$の近くで制限しない，条件が無いのと同じ！よって条件無し問題．
こうして場合分けすれば春学期の等式条件か条件無しに帰着．不等式条件が複数ある場合は場合分けの組合せを全部やればよい．
でも，場合分けは煩雑で具体例によって違うので，昔の人がGordanの定理を使って見かけ上場合分けを避ける工夫をした結果がＦＪ条件．

Gordanの定理（教科書p.27第5章定理1）

Gordanの定理. $\overrightarrow{a}_1, \ldots, \overrightarrow{a}_m \in \mathbb{R}^n$**について次の**(1)**と**(2)**は同値：**

(1) $\overrightarrow{p} \geqq \overrightarrow{0}$ **かつ** $\overrightarrow{p} \neq \overrightarrow{0}$ **なる** $\overrightarrow{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix} \in \mathbb{R}^m$

（非負要素たちからなるn**次元ベクトル）であって**
$p_1 \overrightarrow{a}_1 + \cdots + p_m \overrightarrow{a}_m = \overrightarrow{0}$ **を満たすものがある.**

(2) $(\overrightarrow{a}_i, \overrightarrow{y}) < 0,\ i = 1, \ldots, m,$ **を満たす** $\overrightarrow{y} \in \mathbb{R}^n$ **はない.**
ここで$(\cdot, \cdot)$**は内積.**

注. 同値とはベクトル$\overrightarrow{a}_i$**たちの組が与えられたときそれについて，**(1)**が成り立っていれば**(2)**が成り立ち，かつ，**(2)**が成り立っていれば**(1)**が成り立つこと.** (2)**が成り立つことを確かめ（証明し）て**(1)**の**$\overrightarrow{p}$**を用いることを**(2)⇒(1)**と略記する.**
Gordanの定理(2)⇒(1)**と次ページ補題を用いて３ページ前の定理が証明できる.**

下り坂の補題（教科書p.27補題）

教科書は単に補題と呼ぶが，引用のためこの講義内でだけ下り坂の補題と名付ける．

下り坂の補題．ベクトル $\overrightarrow{u} \in \mathbb{R}^n$ が $\nabla f(\overrightarrow{a})\overrightarrow{u} < 0$ を満たせば，

$\exists t_0 > 0;\ (\forall 0 < t < t_0)\ f(\overrightarrow{a} + t\overrightarrow{u}) < f(\overrightarrow{a})$．

注．勾配ベクトルは行ベクトル，位置ベクトルを含むその他のベクトルは列ベクトルとしているので，行列の積の表記 $\nabla f(\overrightarrow{a})\overrightarrow{u}$ は２つのベクトルの内積に等しい．その他記号は春学期講義ｐｄｆ前半を参照．

- 次回は下り坂の補題の証明と，Gordan の定理は認めてそれと下り坂の補題から定理２（FJ条件）の証明の概要．

積み残したことの予告．

- 講義ｐｄｆ１０で分離定理を証明し，それを用いて
- 講義ｐｄｆ１１で Gordan の定理を証明する．
- 講義前半では Gordan の定理が正しいことを認めて定理２の証明に進むので,「証明が済んだ」感じがしないので，分離定理と Gordan の定理の証明を終えた後，講義ｐｄｆ１２で再度下り坂の補題と定理２の証明を数式を用いてやり直す．

FJ条件と Gordan の定理に関係ありそうな理由

冒頭のページのとおり，引き続き次回まで f や g_i たちや $\overrightarrow{a}$ は**冒頭の青字を満たす**として話を進める．

$m=2$, $g_1(\overrightarrow{a})<0$, $g_2(\overrightarrow{a})=0$ の場合．
${}^t\overrightarrow{a}_1=\nabla f(\overrightarrow{a})$ および ${}^t\overrightarrow{a}_2=\nabla g_2(\overrightarrow{a})$ と置く（転置は前ページの注のとおり $\overrightarrow{a}_1$ と $\overrightarrow{a}_2$ は列ベクトル，勾配ベクトルは行ベクトルなので合わせるため）．
$\nabla f(\overrightarrow{a})\overrightarrow{u}<0$ と $\nabla g_2(\overrightarrow{a})\overrightarrow{u}<0$ を同時に満たす $\overrightarrow{u}\in\mathbb{R}^n$ が無いことが分かったとすると，Gordan の定理の (2) が（$\overrightarrow{y}$ を $\overrightarrow{u}$ と書いて）成り立つので，
Gordan の定理の (2) ⇒ (1) から
$p_1\nabla f(\overrightarrow{a})+p_2\nabla g_2(\overrightarrow{a})=\overrightarrow{0}$ を満たす同時に 0 でない $p_1\geqq 0$ と $p_2\geqq 0$ がある．
$\mu_0=p_1$, $\mu_1=0$, $\mu_2=p_2$ と置くと（以下各自確かめよ），
μ_0,μ_1,μ_2 は非負で同時に 0 にならず $\mu_1g_1(\overrightarrow{a})=\mu_2g_2(\overrightarrow{a})=0$ と $\mu_0\nabla f(\overrightarrow{a})+\mu_1\nabla g_1(\overrightarrow{a})+\mu_2\nabla g_2(\overrightarrow{a})=\overrightarrow{0}$ が成り立つ．FJ 条件が得られた！
$\nabla f(\overrightarrow{a})\overrightarrow{u}<0$ と $\nabla g_2(\overrightarrow{a})\overrightarrow{u}<0$ を同時に満たす $\overrightarrow{u}\in\mathbb{R}^n$ が無いことを，**極小値**を取る仮定と**下り坂の補題**から得る．

2．Gordanの定理からFritz–John条件の導出

目次．

設定（前回冒頭から）． m と n は正整数．f，g_1，g_2，...，g_m は n 変数 C^1 級実数値関数たち．$\overrightarrow{a} \in \mathbb{R}^n$ は $g_1(\overrightarrow{a}) \leqq 0$, ..., $g_m(\overrightarrow{a}) \leqq 0$ を満たす点．**以上を１組与えられたとして講義１と２で固定する．**

定理２（Fritz–John条件）． 条件 $\overrightarrow{g} \leqq 0$ の下での極小値を f が $\overrightarrow{a}$ で取るならば，$\overrightarrow{a}$ でのFritz-John (FJ)条件が成り立つ．

今回は定理２の証明： $I(\overrightarrow{a}) = \{i \mid g_i(\overrightarrow{a}) = 0\}$ と置く．

- $g_i(\overrightarrow{a}) < 0$ $(i \notin I(\overrightarrow{a}))$ は制限にならない条件 → $\mu_i = 0$，
- $g_i(\overrightarrow{a}) = 0$ $(i \in I(\overrightarrow{a}))$ は，「$g_i < 0$ 側で f が大きい」 →
  下り坂の補題と極小値を取る仮定から $f(\overrightarrow{a} + t\overrightarrow{u}) < f(\overrightarrow{a})$, $g_i(\overrightarrow{a} + t\overrightarrow{u}) < 0$, $i \in I(\overrightarrow{a})$, $0 < t < t_0$, なる $\overrightarrow{u}$ が無い
  → Gordanの定理(2)⇒(1)が使える

下り坂の補題（講義ｐｄｆ１再掲）

補題（再掲）． $\overrightarrow{u} \in \mathbb{R}^n$が$\nabla f(\overrightarrow{a})\overrightarrow{u} < 0$を満たせば，十分小さな正の$t$に対して$f(\overrightarrow{a} + t\overrightarrow{u}) < f(\overrightarrow{a})$．

注１． $\nabla f(\overrightarrow{a})\overrightarrow{u}$はベクトルの内積の行列表記：

$$\nabla f(\overrightarrow{a})\overrightarrow{u} = u_1 \frac{\partial f}{\partial x_1}(\overrightarrow{a}) + \cdots + u_n \frac{\partial f}{\partial x_n}(\overrightarrow{a}))$$

注２．「十分小さな正のtに対して…」とは，（小さいかも知れないが正の）$t_0 > 0$があって$0 < t < t_0$ならば…，の意味．

気持ち．

- 関数fを地図上の東経北緯で表した各点の標高とすると，$\nabla f(\overrightarrow{a})$は地点$\overrightarrow{a}$での勾配（傾斜）の最も急な登りの方向（前回１の問参照）．
- $\nabla f(\overrightarrow{a})\overrightarrow{u} < 0$はそれと鈍角をなす方向，つまり下り坂の方向．
- 補題：勾配ベクトルと鈍角の方向のfの値はしばらく小さい．

下り坂の補題の証明（の教科書記述水準の）概要

$h(t) = f(\overrightarrow{a} + t\overrightarrow{u})$ で関数 h を定義すると，合成関数の微分法則（教科書第２章定理4(1)，春学期講義ｐｄｆ７）から $h'(0) = \nabla f(\overrightarrow{a})\overrightarrow{u}$ なので補題の仮定から $h'(0) < 0$．

微分 $h'(0)$ は原点 $t = 0$ での関数 h の平均変化率 $\dfrac{h(t) - h(0)}{t}$ の極限だから，$h'(0) < 0$ は $t > 0$ かつ t が小さいとき $h(t) - h(0) < 0$ を意味する．（証明のために一時的に）$h(t) = f(\overrightarrow{a} + t\overrightarrow{u})$ と置いていたので，元の記号に戻すと，$f(\overrightarrow{a} + t\overrightarrow{u}) < f(\overrightarrow{a})$ が小さい正の t に対して成り立つことを意味する（証明終わり）．

式による精密・正確な証明は１年の微分の知識で可能（各自確かめよ）だが，講義ｐｄｆでは最終回に回す．

Gordanの定理の(2)⇒(1)（講義ｐｄｆ１再掲）

Gordanの定理(2)⇒(1)（「(2)ならば(1)」だけ使うので再掲）．
$\overrightarrow{a}_1, \ldots, \overrightarrow{a}_m \in \mathbb{R}^n$について，
内積を$(\cdot, \cdot)$と書くとき，$(\overrightarrow{a}_i, \overrightarrow{y}) < 0,\ i = 1, \ldots, m,$ を満たす$\overrightarrow{y} \in \mathbb{R}^n$が無いことがわかれば，

$$\overrightarrow{p} \geqq \overrightarrow{0}\text{ かつ }\overrightarrow{p} \neq \overrightarrow{0}\text{ を満たす }\overrightarrow{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix} \in \mathbb{R}^m$$

（非負要素からなる零ベクトルでないm次元ベクトル）であって，
$p_1 \overrightarrow{a}_1 + \cdots + p_m \overrightarrow{a}_m = \overrightarrow{0}$ を満たすものがある．

注１．講義ｐｄｆ１では(1)と(2)が同値と書いたが，FJ条件の導出（定理２の証明）に使うのは(2)⇒(1)の向きだけ．
注２．Gordanの定理は分離定理を講義ｐｄｆ１０で証明した後に講義ｐｄｆ１１で分離定理から証明する．今はGordanの定理が正しいことを信じて話を進める．

定理２の証明

- $I(\overrightarrow{a}) = \{i \in \{1,\ldots,m\} \mid g_i(\overrightarrow{a}) = 0\}$ と置く．

$I(\overrightarrow{a})$ は条件の関数 g_i たちの添字 i の集合．

- $\overrightarrow{a} \in \mathbb{R}^n$ は不等式条件 $g_1(\overrightarrow{a}) \leqq 0, \cdots, g_m(\overrightarrow{a}) \leqq 0$ を満たし，g_i たちは C^1 級なので特に連続関数（講義ｐｄｆ１冒頭青字設定）．よって，$i \notin I(\overrightarrow{a})$ ならば $g_i(\overrightarrow{a}) < 0$，$i \in I(\overrightarrow{a})$ ならば $g_i(\overrightarrow{a}) = 0$．
- $g_i(\overrightarrow{a}) < 0$ のとき，連続関数なので $\overrightarrow{a}$ を中心とするある球内（$\overrightarrow{a}$ の近傍）で $g_i(\overrightarrow{x}) < 0$ （春学期講義ｐｄｆ前半参照）．$g_i(\overrightarrow{x}) \leqq 0$ の条件を追加しても削除しても，$\overrightarrow{a}$ で極小値を取るかには（$\overrightarrow{a}$ の近くの点を制限しないので）影響がない．以下条件の一覧から除外する．
- $g_i(\overrightarrow{a}) = 0$ のとき，$g_i(\overrightarrow{x}) > 0$ を満たす $\overrightarrow{x}$ は $\overrightarrow{a}$ のどんなに近くでも（不等式条件を満たさないから）極小の競争に参加しない．「等高線」$g_i = 0$ を挟んで反対側 $g_i(\overrightarrow{x}) \leqq 0$ だけが考慮対象．

（続）定理２の証明（補題の適用）

$\overrightarrow{x} = \overrightarrow{a} + t\overrightarrow{u}$ は t が0に近いと $\overrightarrow{a}$ に近い．前ページで注目した $i \in I(\overrightarrow{a})$ $(g_i(\overrightarrow{a}) = 0)$ を満たす i たちすべてで $g_i(\overrightarrow{a} + t\overrightarrow{u}) < 0$ を満たす点は不等式条件をみたすので，$\overrightarrow{a}$ で f が極小であるには t がある程度0に近いと $f(\overrightarrow{a} + t\overrightarrow{u}) \geqq f(\overrightarrow{a})$ が必要である．対偶で書くと，

- $f(\overrightarrow{a} + t\overrightarrow{u}) < f(\overrightarrow{a})$, $g_i(\overrightarrow{a} + t\overrightarrow{u}) < 0$, $i \in I(\overrightarrow{a})$, すべての不等号が $t \to 0$ で同時には成り立たなくなる．

ここで背理法の仮定として，もし $\nabla f(\overrightarrow{a})\overrightarrow{u} < 0$, $\nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0$, $i \in I(\overrightarrow{a})$, を同時に満たす $\overrightarrow{u}$ があると，下り坂の補題から，上記不等式が同時にいくらでも小さな正の t で成り立つから矛盾．背理法で

- $\nabla f(\overrightarrow{a})\overrightarrow{u} < 0$, $\nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0$, $i \in I(\overrightarrow{a})$, を同時に満たす $\overrightarrow{u} \in \mathbb{R}^n$ はない．

（続）定理２の証明（Gordanの定理の適用）

$g_i(\overrightarrow{a}) = 0$を満たす$i$の個数に$1$を加えた数を$m$と置いて，$m$個の勾配ベクトル$\nabla f(\overrightarrow{a})$, $\nabla g_i(\overrightarrow{a})$, $i \in I(\overrightarrow{a})$, を（教科書によって勾配ベクトルは行ベクトルの扱いなので）転置して列ベクトルにしたものを順に通し番号で$\overrightarrow{a}_1, \ldots, \overrightarrow{a}_m$と置く．さらに$\overrightarrow{y} = \overrightarrow{u}$と置くと，前ページ最後の結論は**Gordanの定理の**(2)（(2)⇒(1)の仮定部分）だから，(1)（(2)⇒(1)の結論）から$\overrightarrow{p} \geqq \overrightarrow{0}$かつ$\overrightarrow{p} \neq \overrightarrow{0}$を満たす$\overrightarrow{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix} \in \mathbb{R}^m$であって，$p_1\overrightarrow{a}_1 + \cdots + p_m\overrightarrow{a}_m = \overrightarrow{0}$を満たすものがある．（元の極値問題の$m$とここの$m$が同じ記号で別の値なのは悪いが，教科書のせい.）

定理２の証明（終わり）

$\overrightarrow{a}_i$たちを極値問題の関数たちに戻すにあたって，$I(\overrightarrow{a})$の要素でない添字の条件を除外したことを思い出す．除外してきた$i \notin I(\overrightarrow{a})$について$\mu_i = 0$と置き，$i \in I(\overrightarrow{a})$が通し番号では$j$（$\overrightarrow{a}_j = \nabla g_i(\overrightarrow{a})$）のとき$\mu_i = p_j$と置く．$\overrightarrow{a}_1 = \nabla f(\overrightarrow{a})$ から通し番号を振ったことを思い出しつつ$\mu_0 = p_1$と置くと先ほどのベクトル方程式は
$\mu_0 \nabla f(\overrightarrow{a}) + \mu_1 \nabla g_1(\overrightarrow{a}) + \cdots + \mu_m \nabla g_m(\overrightarrow{a}) = \overrightarrow{0}$ となる．

前ページの添字の付け直し作業から，元の極値問題の添字iすべてで$g_i(\overrightarrow{a}) = 0$または$\mu_i = 0$．よって$\mu_i g_i(\overrightarrow{a}) = 0,\ i = 1, \ldots, m$．Gordanの定理から$\overrightarrow{p}_i \geqq \overrightarrow{0}$かつ$\overrightarrow{p} \neq \overrightarrow{0}$なので前ページの添字の付け直し作業と合わせて$\mu_i \geqq 0,\ i = 0, \ldots, m$，かつ，$(\mu_0, \ldots, \mu_m) \neq \overrightarrow{0}$．よってFJ条件を得た（証明終わり）．

3．Kuhn–Tucker条件と制約想定

目次．

- Kuhn–Tucker**条件（以下講義内略記で**KT**条件）：**$\exists \lambda_1, \cdots, \lambda_m;$
$\nabla f(\overrightarrow{a}) + \lambda_1 \nabla g_1(\overrightarrow{a}) + \cdots + \lambda_m \nabla g_m(\overrightarrow{a}) = \overrightarrow{0}$,
$\lambda_1 g_1(\overrightarrow{a}) = \cdots = \lambda_m g_m(\overrightarrow{a}) = 0$,
$\lambda_1 \geqq 0, \ldots, \lambda_m \geqq 0$．

- **制約想定：不等式条件下極小値問題から（**FJ**条件より強い）**KT**条件を得るための不等式条件たちについての十分条件．**

- Cottle**の制約想定：**$\exists \overrightarrow{u} \in \mathbb{R}^n;\ \nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0,\ i \in I(\overrightarrow{a})$．

Fritz–John条件とKuhn–Tucker条件

引き続き講義ｐｄｆ１冒頭の設定を満たす関数たち f と g_i たちと点 $\overrightarrow{a}$ を固定：
m と n は正整数．n 変数 C^1 級実数値関数たち f，g_1，g_2，…，g_m．
点 $\overrightarrow{a} \in \mathbb{R}^n$ は不等式条件 $g_1(\overrightarrow{a}) \leqq 0, \ldots, g_m(\overrightarrow{a}) \leqq 0$ を満たす．

- $\overrightarrow{a}$ でＫＴ条件が成り立つとは，$\exists\lambda_1, \cdots, \lambda_m;$

$\nabla f(\overrightarrow{a}) + \lambda_1 \nabla g_1(\overrightarrow{a}) + \cdots + \lambda_m \nabla g_m(\overrightarrow{a}) = \overrightarrow{0}$,
$\lambda_1 g_1(\overrightarrow{a}) = \cdots = \lambda_m g_m(\overrightarrow{a}) = 0$,
$\lambda_1 \geqq 0, \ldots, \lambda_m \geqq 0$．

- $\overrightarrow{a}$ でＦＪ条件が成り立つとは（講義ｐｄｆ１），$\exists\mu_0, \mu_1, \ldots, \mu_m;$

$\mu_0 \nabla f(\overrightarrow{a}) + \mu_1 \nabla g_1(\overrightarrow{a}) + \cdots + \mu_m \nabla g_m(\overrightarrow{a}) = \overrightarrow{0}$,
$\mu_1 g_1(\overrightarrow{a}) = \cdots = \mu_m g_m(\overrightarrow{a}) = 0$,
$\mu_0 \geqq 0, \mu_1 \geqq 0, \ldots, \mu_m \geqq 0$,
$(\mu_0, \mu_1, \cdots, \mu_m) \neq \overrightarrow{0}$．ＫＴ条件と比べると，設定の関数たちについて，

命題． $\overrightarrow{a}$ でＦＪ条件を満たし $\mu_0 \neq 0$ であることと $\overrightarrow{a}$ でＫＴ条件を満たすことは同値である．（明らかだが，証明を各自確かめよ.）

極小値問題とＫＴ条件（教科書p.28第５章定理3）

定理．講義ｐｄｆ１冒頭の設定（前ページ再掲）を満たす関数fとg_iたちと点$\overrightarrow{a}$について，講義ｐｄｆ２の記号$I(\overrightarrow{a}) = \{i \mid g_i(\overrightarrow{a}) = 0\}$に対して$\nabla g_i(\overrightarrow{a})$, $i \in I(\overrightarrow{a})$, が１次独立であって，$f$が$\overrightarrow{a}$で$g_i \leqq 0$, $i = 1, \ldots, m$, の下での極小値を取れば$\overrightarrow{a}$でＫＴ条件が成り立つ．

証明．設定と極小値を取る条件の下でＦＪ条件が成り立つことは講義ｐｄｆ２で見たとおりなので，前ページ最後の命題から，$\mu_0 \neq 0$を証明すればＫＴ条件も成り立つから，以下背理法の仮定で$\mu_0 = 0$とする．ＦＪ条件が成り立つので特にベクトル方程式

$\mu_0 \nabla f(\overrightarrow{a}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0}$ が成り立つ．（次ページに続く）

（続）定理３の証明

（前ページからの続き）$I(\overrightarrow{a}) = \{i \mid g_i(\overrightarrow{a}) = 0\}$ （講義ｐｄｆ２参照）なので$i \notin I(\overrightarrow{a})$ならば$g_i(\overrightarrow{a}) \neq 0$．よってＦＪ条件の$\mu_i g_i(\overrightarrow{a}) = 0$から$i \notin I(\overrightarrow{a})$ならば$\mu_i = 0$．背理法の仮定の$\mu_0 = 0$と合わせてベクトル方程式に代入すると $\displaystyle\sum_{i \in I(\overrightarrow{a})} \mu_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0}$．

また$\mu_i = 0$, $i \notin I(\overrightarrow{a})$，と$\mu_0 = 0$はＦＪ条件の$(\mu_0, \mu_1, \ldots, \mu_m) \neq \overrightarrow{0}$と合わせると$i \in I(\overrightarrow{a})$なる$\mu_i$に０以外があることを意味する．したがって，$\displaystyle\sum_{i \in I(\overrightarrow{a})} \mu_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0}$は$\nabla g_i(\overrightarrow{a})$, $i \in I(\overrightarrow{a})$，が１次従属であることを意味する．これは１次独立という定理の仮定に矛盾するので，背理法から$\mu_0 \neq 0$である（証明終わり）．

KT条件がだいじにされる理由（個人的想像）

$\mu_0 = 0$のときFJ条件に関数fが現れない（各自確かめよ）ので，そのときfに対してFJ条件が成り立てば$-f$に対しても成り立つ．一方が極小なら他方は極大なので両方とも同じ点$\overrightarrow{a}$で極小にはならないから，FJ条件は$\overrightarrow{a}$で極小値を取ることの十分条件になりえない．
他方，関数（適用範囲）を制限すれば，KT条件は（必ずfが入っているので）極小値を取ることの十分条件になりえるし，実際，KT条件は，関数たちが凸関数ならば最小値の十分条件になる．その伏線として教科書でこの段階からKT条件が紹介されている．

制約想定

- **制約想定**：**不等式条件下極小値問題から（FJ条件より強い）KT条件を得るための不等式条件たちについての十分条件．**

ー　$\mu_0 \neq 0$ がその条件であることは命題で指摘済みだが，FJ条件の方程式を解かないといけないから,（解いたらKT条件いらないから）実用上は役に立たない．

ー　定理３から $\nabla g_i(\overrightarrow{a})$, $i \in I(\overrightarrow{a})$, が１次独立であるという条件もあるが，微分の計算，$g_i(\overrightarrow{a}) = 0$ を満たす i の選択，１次独立性の判定の作業が必要．

ー　１次独立性の判定は回避できるのがCottleの制約想定．（凸関数の場合に微分を計算しなくてよいSlaterの制約想定と同値であることの伏線の意味のほうが大きい.）

- **Cottleの制約想定**：講義ｐｄｆ１冒頭の設定（今回の講義２ページ目に再掲）を満たす関数 g_i たちと $\overrightarrow{a}$ と定理２の添字集合 $I(\overrightarrow{a})$（定理３の証明に再掲）について，$\nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0$, $i \in I(\overrightarrow{a})$, を満たす $\overrightarrow{u} \in \mathbb{R}^n$ が存在すること．

Cottleの制約想定（教科書第５章p.29定理4）

定理．講義ｐｄｆ１冒頭の設定（今回の講義に再掲）を満たす関数fとg_iたちと点$\overrightarrow{a}$についてCottleの制約想定が成り立つとき，fが$\overrightarrow{a}$で$g_i \leqq 0,\ i = 1, \ldots, m,$ の下での極小値を取れば$\overrightarrow{a}$でKT条件が成り立つ．

証明．不等式条件下で極小値を取るのでFJ条件が成り立つ．FJ条件とKT条件についての前述の命題から（定理３と無関係に），$\mu_0 \neq 0$を証明すれば良い．背理法の仮定で$\mu_0 = 0$とすると，定理３の証明と同様にFJ条件のベクトル方程式は$\sum_{i \in I(\overrightarrow{a})} \mu_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0}$となる．$\overrightarrow{u}$との内積を取ると行列表記で

$\sum_{i \in I(\overrightarrow{a})} \mu_i (\nabla g_i(\overrightarrow{a}) \overrightarrow{u}) = \overrightarrow{0}\,\overrightarrow{u} = 0$．Cottleの制約想定から内積はすべて負だが，FJ条件からμ_iたちはすべて非負で０でないものがあるから，負の項の和になって右辺０に等しくないから矛盾．背理法から$\mu_0 \neq 0$（証明終わり）．

教科書 p.29 下半分の注

財の量など変数の非負条件も不等式条件として扱えるが，対応するKT条件の係数は（具体例によらず統一的に）消去可能，という注である．
変数のn個の成分x_jたち全て非負の場合を扱う．講義ｐｄｆ１の注意に従って$x_j \geqq 0$は$-x_j \leqq 0$と直して$-x_j$, $j = 1, \dots, n$, をg_iたちの一部とする．その勾配ベクトルは$-{}^t\overrightarrow{e}_j$（第$j$成分だけ1で残りが0の基本単位ベクトル）．対応するKT条件のベクトル方程式の係数をν_jと置くとベクトル方程式の第j成分は

$\nu_j = \dfrac{\partial f}{\partial x_j}(\overrightarrow{a}) + \displaystyle\sum_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_j}(\overrightarrow{a})$．よって$\overrightarrow{\nu} = \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix}$と置くと，係数の非負条件と合わせて$\nabla f(\overrightarrow{a}) + \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(\overrightarrow{a}) = {}^t\overrightarrow{\nu} \geqq \overrightarrow{0}$．　　　———— (*1)

（次ページに続く）

（続）教科書p.29下半分の注

（前ページからの続き）
$\lambda_i g_i(\vec{a})=0$ **の中にも** ν_j **が** $(x_j=a_j$ **として**$)$ $-\nu_j a_j=0$ **の形で入るが，** $\nu_j \geqq 0$ **かつ非負変数条件** $a_j \geqq 0$ **の下では，全て加えた** ${}^t\vec{\nu}\,\vec{a}=\displaystyle\sum_{j=1}^{n}\nu_j a_j=0$ **と各項が** 0 **と**
が同値． $\vec{\nu}$ **に** $(*1)$ **を代入して** $\displaystyle(\nabla f(\vec{a})+\sum_{i=1}^{m}\lambda_i \nabla g_i(\vec{a}))\vec{a}=0$ ——(*2)
よって $(*1)(*2)$ **と，（非負変数条件関係以外の）不等式条件** $g_i\leqq 0$ **から得る**
$\lambda_i g_i(\vec{a})=0$ **と** $\lambda_i \geqq 0,\ i=1,\ldots,m,$
の連立条件が KT **条件と同値である（以上各自確かめよ）．**
注１．極値の候補を与える方程式としては，以上に加えて元の不等式条件 $g_i\ \leqq\ 0,$ $i=1,\ldots,m,$ **と非負変数条件** $x_j \geqq 0,\ j=1,\ldots,n,$ **も加える．**
注２．制約想定等が無いときは，以上で得る解の他に，FJ **条件のうち** $\mu_0=0$ **の場合の解が候補として残る．**

４．例１（教科書 p.30–31）

目次．

$-x_1^2 + 2x_2 \leqq 0,\ x_1^2 + x_2^2 - 3 \leqq 0$ **の条件下で** $x_1 + x_2$ **の最大値．**

今回と次回は教科書の具体例をやってみる．

作業手順要約

以下 f や g_i の関数名や点 $\overrightarrow{a}$ などは講義ｐｄｆ１冒頭青字の設定に準拠する．

(i) **最大値問題（例：効用や利益）**は，最大値を求めたい関数 U に対して $f=-U$ と置いてから教科書の定理を使う．不等式条件も $h_1 \geqq 0$（例：非負変数条件）は $g_1=-h_1$，$h_2 \leqq h_3$,（例：所得制約）は $g_2=h_2-h_3$ として，すべて $g_i \leqq 0$ の形に直す．

注．最大値でＦＪ条件を書いたり不等式の向きが違う教科書に準拠しても良いが，一度決めたら生涯その教科書と同じ向きの本しか使わない覚悟で．

(ii) $S=\{\overrightarrow{x} \in \mathbb{R}^n \mid g_i(\overrightarrow{x}) \leqq 0,\ i=1,\ldots,m\}$ （条件を満たす変数の集合（高校の用語では領域））が**有界閉集合ならば最大値の原理**（**春学期**講義ｐｄｆ参照）から，ＦＪ条件で得る極小値の候補のどれかが最小値（**極小値問題だけでなく最小値**問題に解があるということ）．

注．教科書の例は，数学として最小限の煩雑さで最小値の存在が保証されるようにこの手順がある．経済学を含む応用で「まっとう」な数学の使い方をする場合は，欲しい（重要な）解が領域の境界ではなく「真ん中」（内点）になるように変数や関数を適切に選ぶはずなので，極小値を得れば十分な場合もあり得る．

（続）作業手順要約

（前ページからの続き）

(iii) FJ 条件または KT 条件で得る連立方程式・不等式は**どう解いても良い**が，迷ったら $\mu_i = 0$ と $\mu_i > 0$ の場合分けで解くことが多い．

注．高校で最小値は区間の内部の極値と区間の端点での値の比較と習う．FJ 条件や KT 条件は不等式条件から生じる集合 S の境界（1 変数なら端点）と内点での極値をまとめて（ラグランジュ関数の条件無し極値のように）扱うところが先人の工夫．解くときは結局場合分けすることが多い．

(iv) 解（極値の候補）が複数あれば f の値の比較で最小値を決定．

(v) 最初に $f = -U$ と置いたときは最小値の符号を逆にして解答．

注．定理は極値を取る必要条件なので『（特定の点）$\overrightarrow{a}$ で極小値を取れば』とあるが，使うとき連立方程式・不等式系として解くので変数 $\overrightarrow{x}$ のまま（教科書は）扱う．

例1の解（その1）

$-x_1^2+2x_2 \leqq 0$, $x_1^2+x_2^2-3 \leqq 0$ の条件下でx_1+x_2の最大値．
最大値を求めるので逆符号にして$f(x,y)=-x-y$で$f:\mathbb{R}^2\to\mathbb{R}$を定義．$g_1(x,y)=-x^2+2y$と$g_2(x,y)=x^2+y^2-3$で$i=1,2$に対して$g_i:\ \mathbb{R}^2\to\mathbb{R}$, を定義すると，不等式条件は$g_i\leqq 0$, $i=1,2$．
不等式条件を満たす集合は$S=\{(x,y)\in\mathbb{R}^2 \mid y\leqq x^2,\ x^2+y^2\leqq 3\}$で，半径$\sqrt{3}$の円内だから有界で，連続関数の等号付き不等号で定義されているから閉集合（いずれも春講義前半参照）．最大値の原理から最小値がある．

$\nabla f(x,y)=(-1\ \ -1)$, $\nabla g_1(x,y)=(-2x\ \ 2)$,
$\nabla g_2(x,y)=(2x\ \ 2y)$だから定理2（FJ条件）から，

$-\mu_0-2\mu_1 x+2\mu_2 x=0$, $-\mu_0+2\mu_1+2\mu_2 y=0$,
$\mu_1(-x^2+2y)=\mu_2(x^2+y^2-3)=0$,
$\mu_i\geqq 0$, $i=0,1,2$, $(\mu_0,\mu_1,\mu_2)\neq(0,0,0)$．
これに加えて$g_i\leqq 0$から$2y\leqq x^2$と$x^2+y^2\leqq 3$．

例１の解（その２）

[1] $\mu_0 = 0$のとき．最初の式から$\mu_1 = \mu_2$または$x = 0$．残りのFJ条件は$\mu_1 + \mu_2 y = 0$, $\mu_1(-x^2 + 2y) = \mu_2(x^2 + y^2 - 3) = 0$, $\mu_i \geqq 0$, $i = 1, 2$, $(\mu_1, \mu_2) \neq (0, 0)$．

[1–i] $\mu_1 = \mu_2$のとき．$\mu_1 = \mu_2 > 0$から残りの条件のうち$\mu_1 + \mu_2 y = 0$と$\mu_1(-x^2 + 2y) = 0$を同時に満たす解は無い．

[1–i] $x = 0$, $\mu_1 \neq \mu_2$のとき．

[1–i–a] $y = 0$のとき．残りの条件のうち$\mu_1 + \mu_2 y = 0$と$\mu_2(x^2 + y^2 - 3) = 0$から$\mu_1 = \mu_2 = 0$となって$(\mu_1, \mu_2) \neq (0, 0)$に反する．

[1–i–b] $y \neq 0$のとき．$\mu_1 y = \mu_1(-x^2 + y) = 0$から$\mu_1 = 0$なので$\mu_1 + \mu_2 y = 0$から$\mu_2 = 0$となって$(\mu_1, \mu_2) \neq (0, 0)$に反する．

例１の解（その３）

[2] $\mu_0 > 0$ のとき．$\lambda_i = \frac{\mu_i}{\mu_0}$ と置くことでＫＴ条件を得る：
$-2\lambda_1 x + 2\lambda_2 x = 1,\ 2\lambda_1 + 2\lambda_2 y = 1,$
$\lambda_1(-x^2 + 2y) = \lambda_2(x^2 + y^2 - 3) = 0,$
$\lambda_i \geqq 0,\ i = 1, 2,$ 以上に加えて $g_i \leqq 0$ から $2y \leqq x^2$ と $x^2 + y^2 \leqq 3$．
[2–i] $\lambda_i > 0,\ i = 1, 2,$ のとき．[2] のＫＴ条件の２行目から $(x, y) = (\pm\sqrt{2}, 1)$．以下すべてのＫＴ条件を満たすから極値の候補．
[2–ii] $\lambda_1 > 0,\ \lambda_2 = 0$ のとき．ＫＴ条件最初の２行から $(x, y) = (-1, \frac{1}{2})$ と $\lambda_1 = \frac{1}{2}$ を得て，すべての条件を満たすから極値の候補．
[2–iii] $\lambda_1 = 0,\ \lambda_2 > 0$ のとき．ＫＴ条件最初の２行から $(x, y) = \frac{\sqrt{6}}{2}(1, 1)$ だが，$g_1(x, y) = \sqrt{6} - \frac{3}{2} > 0.5 > 0$ で条件を満たさない．
[2–iv] $\lambda_1 = \lambda_2 = 0$ のとき．[2] のＫＴ条件の１行目を満たさない．

例１の解（その４）

以上から極値を取る点の候補は$(x,y)=(\pm\sqrt{2},1),(-1,\frac{1}{2})$．
$f(x,y)=-x-y$の値はそれぞれ
$f(\pm\sqrt{2},1)=\mp\sqrt{2}-1,\ f(-1,\frac{1}{2})=\frac{1}{2}$だから，
最小値は$f(\sqrt{2},1)=-\sqrt{2}-1$．
元の問題は$x+y$の最大値だったことを思い出すと
点$(\sqrt{2},1)$で最大値$\sqrt{2}+1$を取る．

5．例２（教科書 p.31–32）

目次．

今回は３変数以上の例．

例２ $x_1^2 + x_2^2 + x_3^2 - 1 \leqq 0$ **の条件下で** $x_1^3 + x_2^3 + x_3^3$ **の最小値．**

例２の解（その１）

$x_1^2+x_2^2+x_3^2-1\leqq 0$ の条件下で $x_1^3+x_2^3+x_3^3$ の最小値．
$f(x,y,z)=x^3+y^3+z^3$ と $g(x,y,z)=x^2+y^2+z^2-1$ で $f:\ \mathbb{R}^3\to\mathbb{R}$ と $g:\ \mathbb{R}^3\to\mathbb{R}$ を定義すると，$g\leqq 0$ の下で f の最小値を求める問題．
不等式条件を満たす集合は $S=\{(x,y)\in\mathbb{R}^2\mid x^2+y^2+z^2\leqq 1\}$ で，半径１の球内だから有界で，連続関数の等号付き不等号で定義されているから閉集合．最大値の原理から最小値がある．
$\nabla f(x,y,z)=(3x^2\ 3y^2\ 3z^2)$ と $\nabla g(x,y,z)=(2x\ 2y\ 2z)$ から定理２のFJ条件は
$3x^2\mu_0+2x\mu_1=0,\ 3y^2\mu_0+2y\mu_1=0,\ 3z^2\mu_0+2z\mu_1=0,$
$\mu_1(x^2+y^2+z^2-1)=0,\ \mu_0\geqq 0,\ \mu_1\geqq 0,\ (\mu_0,\mu_1)\neq(0,0)$．
これに加えて $g\leqq 0$ から $x^2+y^2+z^2\leqq 1$．

例２の解（その２）

[0] $\mu_0 = 0$のとき．$\mu_1 \neq 0$だから，FJ条件の１行目から$(x, y, z) = (0, 0, 0)$だが$\mu_1(x^2 + y^2 + z^2 - 1) = 0$が成り立たない．

[1] $\mu_0 > 0$のとき．$\lambda = \dfrac{\mu_1}{\mu_0}$と置けばKT条件を得る：

$x(3x + 2\lambda) = y(3y + 2\lambda) = z(3z + 2\lambda) = 0,$

$\lambda(x^2 + y^2 + z^2 - 1) = 0,\ \lambda \geqq 0$．

これに加えて$x^2 + y^2 + z^2 \leqq 1$．

KT条件の１行目から各変数x, y, zそれぞれについて０か$-\frac{2}{3}\lambda$．

方程式不等式系がx, y, zの入れ替えで形を変えないので，ある(x, y, z)が解なら，その順序を入れ替えた組も解になる．よって，変数のうち０の個数で分類できる．

例２の解（その３）

[1–i] $(x, y, z) = (0, 0, 0)$ のとき．KT条件の2行目から $\lambda = 0$．条件をすべて満たすので極小値を取る点の候補．

$\lambda = 0$ なら $(x, y, z) = (0, 0, 0)$ なので，以下 $\lambda > 0$ としてよい．

このときKT条件2行目から $x^2 + y^2 + z^2 = 1$． ———— (*)

[1–ii] $\lambda > 0$ かつ $(x, y, z) = (-\frac{2}{3}\lambda, 0, 0), (0, -\frac{2}{3}\lambda, 0), (0, 0, -\frac{2}{3}\lambda)$ のとき．(*) と $\lambda > 0$ から $\lambda = \frac{3}{2}$ で，条件を全て満たすから $(x, y, z) = (-1, 0, 0), (0, -1, 0), (0, 0, -1)$ はいずれも極小値を取る点の候補．

例２の解（その４）

[1–iii] $\lambda > 0$ かつ

$(x, y, z) = -\frac{2}{3}\lambda\,(1, 1, 0), -\frac{2}{3}\lambda\,(1, 0, 1), -\frac{2}{3}\lambda\,(0, 1, 1)$ のとき．

[1–ii] と同様に

$(x, y, z) = -\frac{1}{\sqrt{2}}(1, 1, 0), -\frac{1}{\sqrt{2}}(1, 0, 1), -\frac{1}{\sqrt{2}}(0, 1, 1)$ および

$\lambda = \frac{3}{2\sqrt{2}}$ で，条件を全て満たすからいずれも候補である．

[1–iv] $\lambda > 0$ かつ $(x, y, z) = -\frac{2}{3}\lambda\,(1, 1, 1)$ のとき．

[1–ii][1–iii] と同様に $(x, y, z) = -\frac{1}{\sqrt{3}}(1, 1, 1)$ および

$\lambda = \frac{\sqrt{3}}{2}$ で，条件を全て満たすから候補である．

例２の解（その５）

各候補について $f(x,y,z)=x^3+y^3+z^3$ の値を計算すると，

[1–i] （原点）： $f(0,0,0)=0$．

[1–ii] （座標成分の 2 つが 0 残りが -1 の点）： $f(x,y,z)=-1$．

[1–iii] （座標成分の 1 つが 0 残りが $-\frac{1}{\sqrt{2}}$ の点）：$f(x,y,z)=-\frac{1}{\sqrt{2}}$．

[1–iv] （点 $-\frac{1}{\sqrt{3}}(1,1,1)$）：$f(-\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}})=-\frac{1}{\sqrt{3}}$．

以上を比較すると，最小値は [1–ii] の

$f(-1,0,0)=f(0,-1,0)=f(0,0,-1)=-1$ である．

６．凸集合と凸関数（教科書第６章p.37,38）

目次．

- $S \subset \mathbb{R}^n$が**凸集合**(convex set)：

$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)(\forall 0 < \lambda < 1)\ \lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y} \in S$

- Sが凸集合のとき$f:\ S \to \mathbb{R}$が**凸関数**(convex function)：

$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)(\forall 0 < \lambda < 1)\ f(\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \leqq \lambda f(\overrightarrow{x}) + (1-\lambda) f(\overrightarrow{y})$

- $\overrightarrow{g}:\ \mathbb{R}^n \to \mathbb{R}^m$が凸関数ならば

$\{\overrightarrow{x} \in \mathbb{R}^n \mid \overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}\}$ は凸集合

凸集合の定義

- n 次元空間の集合 $S \subset \mathbb{R}^n$ が**凸集合**とは，

S のどの２点のどの内分点も S の点であること．

- 式で書くと，$S \subset \mathbb{R}^n$ が凸集合とは，

$$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)(\forall 0 < \lambda < 1)\ \lambda \overrightarrow{x} + (1 - \lambda)\overrightarrow{y} \in S$$

記号の復習．($\forall$ は**春学期**経済数学１の講義ｐｄｆ２を復習せよ．ベクトルは**空間 $\mathbb{R}^n$ の点で，点についての演算は数ベクトル（位置ベクトル＝座標表示）としての演算**（経済数学１の講義ｐｄｆ３も参照）．

- 式は（つまり S が凸集合であることは）

$$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)\ \{\lambda \overrightarrow{x} + (1 - \lambda)\overrightarrow{y} \mid 0 \leqq \lambda \leqq 1\} \subset S$$ と同値

- 言葉で書くと，$S \subset \mathbb{R}^n$ が凸集合とは，

S のどの２点を端点とする線分も S の部分集合であること．

- ２つの式は点ごとに見た場合と集合（図形）で見た場合に対応

凸集合の定義（補足など）

- 雑な例：膨らんだ風船は凸集合．餌を食べる際のパックマンは凸集合でない．
- $\lambda = 0$と$\lambda = 1$（$\overrightarrow{x}$と$\overrightarrow{y}$）はSの要素の前提なので改めて確かめる必要が無いから定義は$0 < \lambda < 1$ （定義が$0 \leqq \lambda \leqq 1$であっても同値）．
- 開集合・閉集合と凸集合は直接関係ない（例：開球も閉球も凸集合）．部分的に境界を含む場合もありうるが注意必要．例：正方形は境界を全部含んでも，全部含まなくても，2辺を含んでも凸集合だが，正方形の内部と4頂点からなる集合は凸集合では無い（各自証明せよ）．
- 講義はいわゆる図形（$\mathbb{R}^n$の凸集合）を扱うが，一般化の際も，定義で内分を用いるから線形結合が定義されること（線形空間）は必要．
- 分離定理はまさに凸集合の性質なので，理論系の経済学では基礎中の基礎概念の1つ（講義終盤，教科書第7章でも扱う）.（経済現象は多様なので，凸性くらいしか一般論で仮定できない事情もありそう.）

凸集合の例

例1．n 次元空間の中の線分

$$S = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid 0 \leqq x_1 \leqq 1,\ x_2 = x_3 = \cdots = x_n = 0 \}$$

略証． $a, b \in [0, 1]$ と $0 < \lambda < 1$ に対して，
$0 = 0 \times \lambda + 0 \times (1 - \lambda) \leqq a\lambda + b(1 - \lambda) \leqq 1 \times \lambda + 1 \times (1 - \lambda) = 1$ だから，

$$\lambda \begin{pmatrix} a \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} b \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a\lambda + b(1 - \lambda) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in S$$

例2．n 次元球 $S = \{ \overrightarrow{x} \in \mathbb{R}^n \mid \|\overrightarrow{x}\| \leqq 1 \}$ **（講義6の問(a)）**

例3．原点を通り $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ **を法線ベクトルとする平面で仕切られた法線ベクトルと逆向き側の半空間** $S = \{ \overrightarrow{x} \in \mathbb{R}^3 \mid (1\ \ 1\ \ 1)\overrightarrow{x} \leqq 0 \}$ **（講義6の問(b)で** $n = 3,\ m = 1$ **の例）**

凸集合の例その２（分離定理で使う集合）

例4．「第３象限」$T = \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid x_i < 0,\ i = 1, \ldots, n\}$

注．もちろん$n = 2$（平面）のとき第３象限と呼ぶ意味で，第$1, 3, 4$象限も凸集合．

例5．mを自然数，Bを$n \times m$行列とするとき，

線形部分空間$V = \{B\overrightarrow{y} \mid \overrightarrow{y} \in \mathbb{R}^m\}$

注．部分空間Vが与えられたとき，その基底$\overrightarrow{e}_j,\ j = 1, \ldots, \mathrm{dim}V$，を横に並べた行列を$B$と置けば$m = \mathrm{dim}V$として上記のように書ける（線形代数続論）．

例6．mを自然数，Bを$n \times m$行列，$\overrightarrow{b} \in \mathbb{R}^n$とするとき，

アフィン空間$A = \{B\overrightarrow{y} + \overrightarrow{b} \mid \overrightarrow{y} \in \mathbb{R}^m\}$

注．部分空間を平行移動した集合．比例$y = kx$と１次式$y = ax + b$の図形の関係．

（以上が凸集合なのは明らかだが定義式を満たすことを各自確かめよ．）

凸関数の定義

- $S \subset \mathbb{R}^n$ が凸集合のとき，$f:\ S \to \mathbb{R}$ が凸関数とは，S のどの２点のどの内分点での f の値も２点での値の内分値以下であること．
- 式で書くと，$S \subset \mathbb{R}^n$ が凸集合のとき，$f:\ S \to \mathbb{R}$ が凸関数とは，

$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)(\forall 0 < \lambda < 1)$

$f(\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \leqq \lambda f(\overrightarrow{x}) + (1-\lambda) f(\overrightarrow{y})$

- 例：$f(x) = x^2$ で定義される $f:\ [-1,1] \to \mathbb{R}$ （各自確かめよ）
- 教科書は不等式条件が複数ある場合に束ねて行数を節約することを考えて，ベクトル値関数 $\overrightarrow{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix}:\ \mathbb{R}^n \to \mathbb{R}^m$ で書くことが多いので…，
- $S \subset \mathbb{R}^n$ が凸集合のとき，$\overrightarrow{g}:\ S \to \mathbb{R}$ が凸関数とは，各成分 $g_i:\ \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, m,$ が凸関数のこと
- まったく束ねて書いて節約しただけなので，これは慌てないためだけ．

凸関数の定義（補足など）

定義（再掲）．$S \subset \mathbb{R}^n$ が凸集合のとき，$f: S \to \mathbb{R}$ が凸関数とは，S のどの２点のどの内分点での f の値も２点での値の内分値以下であること．

- S は f の定義域（f の値の存在を保証すべき点の集合）．それを凸集合に限る，という注意が定義冒頭の気持ち．
- その理由（分かれば当たり前）：$\overrightarrow{x}$ と $\overrightarrow{y}$ での凸関数の値がわかる（定義域に入っている）とき，そのどの内分点での値も値 $f(\overrightarrow{x})$ と $f(\overrightarrow{x})$ の内分値以下，と言えないといけないから，f が凸関数と言うには $\overrightarrow{x}$ と $\overrightarrow{y}$ のすべての内分点（＝線分 $\overline{\overrightarrow{x}\,\overrightarrow{y}}$）での f の値が必要．よって定義域は凸集合の必要．

- 凸関数は，グラフでは下向きに凸（例：$f(x) = x^2$）．下向きに凸なのを基準にするのは多分２階微分の符号が正だから．
- f が凹関数とは $-f$ が凸関数のこと
- 効用関数は凹（経済学では違和感あるかも知れないが，２階微分の符号を考えると，この定義は覆らないので，生涯これで我慢すること）

凸関数が定める凸集合（教科書p.37第6章定理1）

定理． m と n を正の整数とする．$\overrightarrow{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^m$ が凸関数ならば

$S^o = \{\overrightarrow{x} \in \mathbb{R}^n \mid g_i(\overrightarrow{x}) < 0, \ i = 1, \ldots, m\}$ および

$\overline{S} = \{\overrightarrow{x} \in \mathbb{R}^n \mid \overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}\}$ は凸集合．

証明． $\overrightarrow{x}, \overrightarrow{y} \in S^o$ および $0 < \lambda < 1$ とすると，凸関数の定義から

$$\begin{aligned} g_i(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y}) &\leqq \lambda g_i(\overrightarrow{x}) + (1-\lambda) g_i(\overrightarrow{y}) \\ &< \lambda \times 0 + (1-\lambda) \times 0 = 0 \end{aligned}$$

が $i = 1, \ldots, m$ について成り立つから S^o の定義と見比べると
$\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y} \in S^o$．
緑字の部分を凸集合の定義と見比べると，S^o が凸集合であることが証明できている．
$\overline{S}$ についての証明もまったく同様（各自証明せよ）（証明終わり）．

注． ベクトルの等号付き不等号はどの教科書でも概ね各成分毎の等号付き不等号なので容認するが，等号無し不等号は教科書によって意味が違うので，使わないことを薦める．講義では使わない．

７．凸関数の微分による特徴付け

目次．

・　**開凸集合** $S \subset \mathbb{R}^n$ **と** $f:\ S \to \mathbb{R}$ **において**

(1) $f \in C^1$ **のとき，**

f **が凸** $\Leftrightarrow$ $f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a})$, $\overrightarrow{x}, \overrightarrow{a} \in S$

(2) $f \in C^2$ **のとき，**

f **が凸** $\Leftrightarrow$ **ヘッセ行列** H_f **が** S **で非負定値**

凸関数と微分（教科書第6章定理2，p.38,39）

定理． $S \subset \mathbb{R}^n$ が開凸集合（開集合かつ凸集合）とする．S で定義された関数 $f: S \to \mathbb{R}$ について以下が成り立つ：
(1) $f \in C^1$ のとき，
f が凸 $\Leftrightarrow$ $f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a})$, $\overrightarrow{x} \in S$, $\overrightarrow{a} \in S$．
(2) $f \in C^2$ のとき，
f が凸 $\Leftrightarrow$ ヘッセ行列 H_f が S 上で非負定値（$H_f(\overrightarrow{x}) \geqq 0$, $\overrightarrow{x} \in S$)．
例． $n = 1$（1変数関数）のときヘッセ行列は2階微分．$f(x) = x^2$ のとき $f''(x) = 2 > 0$, $x \in \mathbb{R}$．

注． (1) の結論の式は $f(\overrightarrow{x}) \geqq f(\overrightarrow{a}) + \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a})$ と書くと，右辺は $\overrightarrow{a}$ での f の接超平面（$n = 1$ なら接線，$n = 2$ なら接平面）．凸関数は，接超平面のグラフが f のグラフよりも下にある．
例． $f(x) = x^2$ の点 a での接線のグラフは $y = 2a(x - a) + a^2$．
$x^2 - (2a(x - a) + a^2) = (x - a)^2 \geqq 0$．

定理の(1)の凸⇒接超平面がfより低いことの証明

講義ｐｄｆ６再掲．fが凸関数とは$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)(\forall 0 < \lambda < 1)$
$f(\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \leqq \lambda f(\overrightarrow{x}) + (1-\lambda) f(\overrightarrow{y})$

まずfが凸関数とすると上記定義から$0 < \lambda < 1$ならば（移項して）
$\lambda(f(\overrightarrow{x}) - f(\overrightarrow{a})) \geqq f(\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{a}) - f(\overrightarrow{a})$．
平均値の定理（春学期）を右辺に用いると$0 < \theta < 1$が存在して，
$f(\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{a}) - f(\overrightarrow{a}) = \nabla f(\overrightarrow{a} + \theta\lambda(\overrightarrow{x} - \overrightarrow{a}))\,\lambda(\overrightarrow{x} - \overrightarrow{a})$．
２つの式をつないで，両辺を$\lambda > 0$で割ると
$f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a} + \theta\lambda(\overrightarrow{x} - \overrightarrow{a}))\,(\overrightarrow{x} - \overrightarrow{a})$．
fはC^1なのでC^0（連続）だから，$\lambda > 0$はいくら0に近くても成り立つためには
$\lambda = 0$での値を用いて
$f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})\,(\overrightarrow{x} - \overrightarrow{a})$．

定理の(1)の接超平面がfより低い⇒凸の証明

逆に$f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a})$ が各 $\overrightarrow{x} \in S$ と $\overrightarrow{a} \in S$ に対して成り立つとする．

$\overrightarrow{x} \in S$, $\overrightarrow{y} \in S$, $0 < \lambda < 1$ に対して，今の仮定で $\overrightarrow{a} = \lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y}$ と置くと，

$f(\overrightarrow{x}) - f(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \geqq \nabla f(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y})(1-\lambda)(\overrightarrow{x} - \overrightarrow{y})$．

$\overrightarrow{a}$ の中身を変えずに $\overrightarrow{x}$ を $\overrightarrow{y}$ に選ぶと

$f(\overrightarrow{y}) - f(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \geqq \nabla f(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y})\lambda(\overrightarrow{y} - \overrightarrow{x})$．

それぞれλ倍と$1-\lambda$倍して加えると

$\lambda f(\overrightarrow{x}) + (1-\lambda)f(\overrightarrow{y}) - f(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y}) \geqq 0$．

移項すればfが凸関数の定義を満たしていることがわかる．

教科書p.39の下のほうの注について

注．春学期の多変数関数の平均値の定理やテイラーの定理（教科書第２章定理7,8）は,「$\overrightarrow{a}$の近傍で成り立つ」となっていたが，$\overrightarrow{a}$と$\overrightarrow{x}$を結ぶ線分を含む開集合でfがC^1ならば，これらの定理の結論が成り立つ.（その形で今回使っている.）
というのは，教科書第２章定理7,8の証明で使った１変数関数の平均値の定理やテイラーの定理は，区間$t \in [0,1]$の各点tで$h(t)$が微分可能ならば，区間$t \in [0,1]$ で$h(1) = h(0) + h'(\theta)$が成り立つ（１変数関数の時は近傍を使わないで最大値の原理を使って証明する）ので，θの近傍で合成関数$h(t) = f(\overrightarrow{a} + t(\overrightarrow{x} - \overrightarrow{a}))$の微分公式を使えば第２章定理7,8の結論の公式を得るからである.
ということを，ここの教科書の注は言っている（春学期にそう書くべきだった）.

定理の(2)のヘッセ行列非負値⇒凸の証明

C^2級なので テイラーの定理が $\overrightarrow{x}, \overrightarrow{a} \in S$ に対して成り立つ：

$$f(\overrightarrow{x}) = f(\overrightarrow{a}) + \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a}) + \frac{1}{2}{}^t(\overrightarrow{x} - \overrightarrow{a})H_f(\overrightarrow{a} + \theta(\overrightarrow{x} - \overrightarrow{a}))(\overrightarrow{x} - \overrightarrow{a})$$

を満たす $0 < \theta < 1$ が存在する．
H_f が非負定値ならば非負定値の定義（春学期の線形代数続論の復習参照）から最後の項は非負だから，$f(\overrightarrow{x}) \geqq f(\overrightarrow{a}) + \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a})$ が $\overrightarrow{x}, \overrightarrow{a} \in S$ に対して成り立つが，これは(1)で証明済みのとおり，f が凸関数であることを意味する．

定理の(2)の凸⇒ヘッセ行列非負値の証明

逆に f が凸関数とすると，証明済みの(1)によって，$\overrightarrow{x} \in S$ と $\overrightarrow{a} \in S$ に対して
$f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})\,(\overrightarrow{x} - \overrightarrow{a})$
が成り立つ．凸集合 S 上で C^2 級なのでテイラーの定理が成り立っているから見比べると，$\frac{1}{2}{}^t(\overrightarrow{x} - \overrightarrow{a})H_f(\overrightarrow{a} + \theta(\overrightarrow{x} - \overrightarrow{a}))(\overrightarrow{x} - \overrightarrow{a}) \geqq 0$ が任意の $\overrightarrow{x}, \overrightarrow{a} \in S$ に対して成り立つ．
（H_f が非負定値と証明したいので,）n 次元ベクトル $\overrightarrow{u} \in \mathbb{R}^n$ と正数 $\epsilon > 0$ を取ってくる．S は凸というだけでなく**開集合**であることも仮定しているので，$\overrightarrow{a} \in S$ の近傍は S に包含されるから，ϵ を十分小さく選べば（つまり $\epsilon_0 > 0$ があって $0 < \epsilon < \epsilon_0$ ならば必ず）$\overrightarrow{x} = \overrightarrow{a} + \epsilon\overrightarrow{u} \in S$ となる．
これを先ほどの不等式に代入して $\frac{1}{2}\epsilon^2 > 0$ で割ると，
${}^t\overrightarrow{u}H_f(\overrightarrow{a} + \theta\epsilon\overrightarrow{u})\overrightarrow{u} \geqq 0,\ 0 < \epsilon < \epsilon_0$．
ϵ がどれほど 0 に近くても成り立ち，f が C^2 級の仮定から H_f は C^0 級（連続）なので，$\epsilon = 0$ と置いた不等式も成り立つ：${}^t\overrightarrow{u}H_f(\overrightarrow{a})\overrightarrow{u} \geqq 0$,
これが任意の $\overrightarrow{u} \in \mathbb{R}^n$ で成り立つから $H_f(\overrightarrow{a})$ は非負定値．
それが任意の $\overrightarrow{a} \in S$ で成り立つから H_f は S 上非負定値である（証明終わり）．

擬凸

- 次ページの７の問に凸関数の基礎的な例を挙げたので，必ず手を動かして確認すること．

- C^1 級関数 $f:\ S \to \mathbb{R}$ が**擬凸**とは
「$\nabla f(\overrightarrow{a})(\overrightarrow{x}-\overrightarrow{a}) \geqq 0$ ならば $f(\overrightarrow{x}) \geqq f(\overrightarrow{a})$」
が成り立つことを言う．
定理の (1) から C^1 級凸関数は擬凸である．(この講義では使わない.)

8．凸関数の最小値の同値条件

目次.

開凸集合 $S \subset \mathbb{R}^n$ と C^1 級の凸関数 $f:\ S \to \mathbb{R}$ について，

- 拘束条件無い場合：

f が $\overrightarrow{a} \in S$ で最小値を取る $\Leftrightarrow\ \nabla f(\overrightarrow{a}) = \overrightarrow{0}$

- 線形等式条件の場合：

$\overrightarrow{b} \in \mathbb{R}^m$ と $m \times n$ 行列 A; $\mathrm{rank}\, A = m < n$ について，

$A\overrightarrow{x} = \overrightarrow{b}$ の下で f が $\overrightarrow{a} \in S$ で最小値を取る

$\Leftrightarrow\ \exists \overrightarrow{\lambda} \in \mathbb{R}^m;\ \nabla f(\overrightarrow{a}) + {}^t\overrightarrow{\lambda} A = \overrightarrow{0}$

凸関数の条件無し極小値問題（教科書第6章定理3）

設定．**開凸集合** $S \subset \mathbb{R}^n$ **を定義域とする** C^1 **級の凸関数** $f: S \to \mathbb{R}$ と $\overrightarrow{a} \in S$ をこの講義で固定する．

まず，条件無し極値問題について，凸関数に限ると追加される結論．
定理．設定の S, f, $\overrightarrow{a}$ について，f が $\overrightarrow{a}$ で S における最小値を取ることと $\nabla f(\overrightarrow{a}) = \overrightarrow{0}$ が成り立つことが同値．
「最小値⇒勾配ベクトル $\overrightarrow{0}$」の証明．$\overrightarrow{a}$ で最小値を取るとする．定義域が開集合ということは境界の点は除いているので，境界で最小値を取ることは設定上できないので内点で最小値を取るから極小値．$\overrightarrow{a}$ で極値を取るから**春学期第３章定理４**から $\nabla f(\overrightarrow{a}) = \overrightarrow{0}$（終わり）．
注．今の部分は凸は関係なく既に分かっていた一般論のまま．**凸のとき付け加わるのは，逆も成り立つこと．**

凸関数の勾配ベクトルが $\overrightarrow{0}$ の点についての証明

例． $f(x)=2x^2-x^4$ の極値を与える点は $f'(x)=0$ の解 $x=0,\pm1$．極小も極大もとり，最大値はあるが開区間では最小値は無い．

例． 凸関数の例で $f(x)=x^2$ だと極小値がただ 1 つでそこで最小値．多変数でもこれが起きる．

「勾配ベクトル $\overrightarrow{0}$ ⇒最小値」の証明． f が凸関数なので，講義ｐｄｆ７の定理の (1) から

$f(\overrightarrow{x})-f(\overrightarrow{a})\geqq\nabla f(\overrightarrow{a})(\overrightarrow{x}-\overrightarrow{a}),\ \overrightarrow{x}\in S$．

$\nabla f(\overrightarrow{a})=\overrightarrow{0}$ なので $f(\overrightarrow{x})\geqq f(\overrightarrow{a}),\ \overrightarrow{x}\in S$．

これは $f(\overrightarrow{a})$ が S での f の最小値という言葉の定義そのものである（証明終わり）．

凸関数の１次式等式条件問題（教科書第６章定理４）

次に（ランクが落ちない）１次等式条件付き極値問題の場合.「１次式＋ランクが落ちない条件」は話を簡単にするため.

定理. 設定の S, f, $\overrightarrow{a}$ について，ベクトル $\overrightarrow{b} \in \mathbb{R}^m$ と $m \times n$ 行列 A で $\mathrm{rank}\, A = m < n$ を満たすものを用いた等式条件 $A\overrightarrow{x} = \overrightarrow{b}$ を $\overrightarrow{x} = \overrightarrow{a}$ が満たすとき，このの条件下での S における最小値を f が $\overrightarrow{a}$ で取ることとと $\nabla f(\overrightarrow{a}) + \overrightarrow{\lambda} A = \overrightarrow{0}$ を満たす $\overrightarrow{\lambda} = (\lambda_1 \ \cdots \ \lambda_m)$（$m$ 個の未定乗数 $\lambda_1, \ldots, \lambda_m$）があることが同値である.

注. 教科書は $\overrightarrow{\lambda}$ を未定乗数を縦に並べた列ベクトルで定義して，ここで $\overrightarrow{\lambda}$ と書いたものを転置 ${}^t\overrightarrow{\lambda}$ と書いていたが，教科書絵勾配ベクトルが行ベクトルであることに注意すると，$\overrightarrow{\lambda} = (\lambda_1 \ \cdots \ \lambda_m)$ と行ベクトルで置くのが自然だからそうする.

最小値を取る場合の証明

「最小値⇒未定乗数法」の証明．開集合で最小値を取るので内点 $\overrightarrow{a}$ での等式条件下の極値だから，春学期第４章定理２の結論が成り立つ．今の等式条件 $A\overrightarrow{x} = \overrightarrow{b}$ は m 本の方程式で第４章定理２との対応は $\overrightarrow{g}(\overrightarrow{x}) = A\overrightarrow{x} - \overrightarrow{b}$ だから $\nabla\overrightarrow{g}(\overrightarrow{x}) = A$ （７の問(c)と講義ｐｄｆ８の答参照）．今の定理の仮定で $\operatorname{rank} A = m$，すなわち，第4章定理２の結論の(1)は起きない．よって結論の(2)（未定乗数法）が起きる．第４章定理２の結論の(2)は

$\nabla f(\overrightarrow{a}) + \overrightarrow{\lambda}\nabla\overrightarrow{g}(\overrightarrow{a}) = \overrightarrow{0}$ と書ける（各自確かめよ）ので，

$\nabla f(\overrightarrow{a}) + \overrightarrow{\lambda}A = \overrightarrow{0}$ となって主張を得る．

注．条件無しの場合と同様に，最小値があるとした場合は凸性は無関係で，春既習の一般論の結果そのままである．

未定乗数法が成り立つ点についての証明

「未定乗数法⇒最小値」の証明．逆に，$\nabla f(\overrightarrow{a}) + \overrightarrow{\lambda} A = \overrightarrow{0}$ が成り立つとすると，f が凸関数なので，講義ｐｄｆ７の定理の (1) から $f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a}),\ \overrightarrow{x} \in S$（条件無しのときと同じ）．

等式条件 $A\overrightarrow{x} = \overrightarrow{b}$ が成り立つとき，$\overrightarrow{a}$ も等式条件が成り立つので，$\nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a}) = -\overrightarrow{\lambda}(A\overrightarrow{x} - A\overrightarrow{a}) = -\overrightarrow{\lambda}(\overrightarrow{b} - \overrightarrow{b}) = \overrightarrow{0}$．よって $\overrightarrow{x} \in S$ かつ $A\overrightarrow{x} = \overrightarrow{b}$ ならば $f(\overrightarrow{x}) \geqq f(\overrightarrow{a})$ だから $\overrightarrow{a}$ で条件付き最小値を取る（証明終わり）．

例（3次元空間の点と平面の距離）

$(a\ b\ c) \neq (0\ 0\ 0)$ とする（教科書で抜けている）．
平面 $ax+by+cz=d$ 上の点 $P(x,y,z)$ と $Q(x_0,y_0,z_0)$ の距離．
距離の代わりに距離の2乗で比較しても同値なので，
$f(x,y,z)=(x-x_0)^2+(y-y_0)^2+(z-z_0)^2$ で $f:\ \mathbb{R}^3 \to \mathbb{R}$ を定義して，f の最小値問題を考える．
$\nabla f(x,y,z)=(2(x-x_0)\ 2(y-y_0)\ 2(z-z_0))$，したがって
$H_f(x,y,z)=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$，だからヘッセ行列 H_f は対角行列で固有値がすべて正（正定値行列）．よって講義ｐｄｆ７の定理の(2) から f は凸関数．
$n=3$（3次元空間上の3変数関数）および $m=1$（平面上という条件）で，平面上の点は $A=(a\ b\ c)$ および $\overrightarrow{b}=d$ と置くと直前の定理の1次式等式条件 $A\overrightarrow{x}=\overrightarrow{b}$ となるので，最小値を取る点は $\nabla f(\overrightarrow{a})+\lambda A=(0\ 0\ 0)$ を満たす.（次ページに続く）

例（続）

（前ページからの続き）これを解くと

$x - x_0 = -\frac{1}{2}\lambda a$, $y - y_0 = -\frac{1}{2}\lambda b$, $z - z_0 = -\frac{1}{2}\lambda c$なので，平面の方程式$ax + by + cz = d$に代入して（$(a\ b\ c) \neq (0\ 0\ 0)$だから$a^2 + b^2 + c^2 > 0$に注意して）$\dfrac{1}{2}\lambda = \dfrac{ax_0 + by_0 + cz_0 - d}{a^2 + b^2 + c^2}$．よって$(x, y, z)$が決まったので$f(x, y, z)$に代入すれば最小値$\dfrac{(ax_0 + by_0 + cz_0 - d)^2}{a^2 + b^2 + c^2}$を得る．距離は$\sqrt{f(x, y, z)}$なので$\dfrac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}$．

９．凸関数の最小値の同値条件（不等式条件）

目次．

- **設定．開凸集合**$S \subset \mathbb{R}^n$**を定義域とする**m**次元ベクトル値**C^1**級凸関数**$\overrightarrow{g} : S \to \mathbb{R}^m$**と**$\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$**を満たす点**$\overrightarrow{a} \in S$**を固定する．**
- **設定の**$\overrightarrow{g}$**が設定の**$\overrightarrow{a}$**で**Slater**の制約想定を満たすとは，**

$\exists \overrightarrow{b} \in S;\ (\forall i \in I(\overrightarrow{a}))\ g_i(\overrightarrow{b}) < 0$

- **定理．設定の**$\overrightarrow{g}$**と**$\overrightarrow{a}$**について，**

Cottle**の制約想定成立と**Slater**の制約想定成立は同値**

- **制約想定は「最小（極小）⇒**KT**条件」のための前提**
- **制約想定と無関係な向き「**KT**条件⇒最小」は凸関数なら成立：**
- **定理．設定の**$\overrightarrow{g}$**と**$\overrightarrow{a}$**について，**C^1**級凸関数**$f : S \to \mathbb{R}$**が**$\overrightarrow{a} \in S$**で**Kuhn–Tucker**条件を満たすならば，**f**は**$\overrightarrow{a}$**で**S**における条件**$\overrightarrow{g} \leqq \overrightarrow{0}$**の下での最小値を取る．**

Slaterの制約想定

- **制約想定**（**講義ｐｄｆ３**）：不等式条件下極小値問題から（FJ条件より強い）KT条件を得るための不等式条件たちについての十分条件．
- **冒頭設定（再掲）**．n, m 正整数，$S \subset \mathbb{R}^n$：開凸，
$\overrightarrow{g}: \ S \to \mathbb{R}^m$：$C^1$ **級凸関数**，$\overrightarrow{a} \in S$：$\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$ を満たす．
- **定義**．設定の $\overrightarrow{g}$ が設定の $\overrightarrow{a}$ でSlater**の制約想定**を満たすとは，
$\exists \overrightarrow{b} \in S; \ (\forall i \in I(\overrightarrow{a})) \ g_i(\overrightarrow{b}) < 0$ を満たすことを言う．
- $I(\overrightarrow{a})$ は**講義ｐｄｆ２**のFJ条件の定理の証明の中の記号：
$i \in I(\overrightarrow{a}) \Leftrightarrow g_i(\overrightarrow{a}) = 0$，$i \notin I(\overrightarrow{a}) \Leftrightarrow g_i(\overrightarrow{a}) < 0$．
- **定理**．**設定の $\overrightarrow{g}$ と $\overrightarrow{a}$ について，**

Cottleの制約想定成立とSlaterの制約想定成立は同値

- Cottleの制約想定：$\exists \overrightarrow{u} \in \mathbb{R}^n; \ \nabla g_i(\overrightarrow{a}) \overrightarrow{u} < 0, \ i \in I(\overrightarrow{a})$.（**講義ｐｄｆ３**）.
- 凸関数に限定したことで，勾配ベクトルを計算する必要のあったCottleの制約想定を元の関数の値だけで済むSlaterの制約想定に同値変形できた．

凸関数の場合の２つの制約想定の同値性の証明

まずSlater⇒Cottleを言う．$\overrightarrow{a}, \overrightarrow{b} \in S$と$\overrightarrow{g}$が$C^1$級凸から，講義ｐｄｆ８の定理(1)の結論を各$i$について（$f = g_i$として）使うと，
$g_i(\overrightarrow{b}) - g_i(\overrightarrow{a}) \geqq \nabla g_i(\overrightarrow{a})(\overrightarrow{b} - \overrightarrow{a})$.
$i \in I(\overrightarrow{a})$のとき，$I(\overrightarrow{a})$の定義から$g_i(\overrightarrow{a}) = 0$かつ定理の仮定から$g_i(\overrightarrow{b}) < 0$なので上の式の左辺は負だから$\overrightarrow{u} = \overrightarrow{b} - \overrightarrow{a}$と置いて$\overrightarrow{u} \in \mathbb{R}^n$を定めると$\nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0,\ i \in I(\overrightarrow{a})$，となってCottleの制約想定を得る．

次にCottle⇒Slaterを言うために，$\nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0,\ i \in I(\overrightarrow{a})$, を満たす$\overrightarrow{u} \in \mathbb{R}^n$があるとする．下り坂の補題（講義ｐｄｆ１）を各$i \in I(\overrightarrow{a})$について（$f = g_i$として）使うと，$i \in I(\overrightarrow{a})$は$g_i(\overrightarrow{a}) = 0$のことだったことも思い出して，

$\exists t_0 > 0;\ (\forall 0 < t < t_0)\ (\forall i \in I(\overrightarrow{a}))\ g_i(\overrightarrow{a} + t\overrightarrow{u}) < g_i(\overrightarrow{a}) = 0$.

たとえば$\overrightarrow{b} = \overrightarrow{a} + \frac{1}{2}t_0\overrightarrow{u}$と選べばSlaterの制約想定が成り立つ．（証明終わり）．

凸とKT条件⇒最小値（教科書p.43第6章定理6）

- 　ＦＪ条件は不等式条件下の極小値の必要条件（講義ｐｄｆ１）だが，関係する関数がすべて凸関数ならばKT条件は十分条件であるという講義ｐｄｆ３の伏線を回収する．
- 　冒頭設定（再々掲）．n, m 正整数，$S \subset \mathbb{R}^n$：開凸，

$\overrightarrow{g}: S \to \mathbb{R}^m$：$C^1$ 級凸関数，$\overrightarrow{a} \in S$：$\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$ を満たす．

定理．冒頭の設定の $\overrightarrow{g}$ と $\overrightarrow{a}$ について，C^1 級凸関数 $f: S \to \mathbb{R}$ が $\overrightarrow{a} \in S$ でKuhn–Tucker条件を満たすならば，f は $\overrightarrow{a}$ で S における条件 $\overrightarrow{g} \leqq \overrightarrow{0}$ の下での最小値を取る．

凸とＫＴ条件⇒最小値の証明

前ページの定理の証明． $\overrightarrow{a}$ で f が最小値を取ることを言いたいので，「$\overrightarrow{x} \in S$ かつ $\overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}$ ならば $f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq 0$」が目標．
f は凸なので

$$f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq \nabla f(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a})$$

は既知（講義ｐｄｆ７定理(1)）．
$\overrightarrow{g}$ も凸なので各成分ごとに同様の不等式が成り立つから，不等号の向きを揃えるべく符号を逆にすると，

$$-\nabla g_i(\overrightarrow{a})(\overrightarrow{x} - \overrightarrow{a}) \geqq -g_i(\overrightarrow{x}) + g_i(\overrightarrow{a}),\ i = 1, \ldots, m.$$

ＫＴ条件が成り立つという仮定から $\lambda_i \geqq 0,\ i = 1, \ldots, m,$ かつ

$$\nabla f(\overrightarrow{a}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0} \text{ と } \lambda_i g_i(\overrightarrow{a}) = 0,\ i = 1, \ldots, m,$$

なので，最初の不等式２行をつなぐことができて（$\lambda_i \geqq 0$ にも注意），

（次ページに続く）

（続）凸とKT条件⇒最小値の証明

（前ページからの続き）

$$f(\overrightarrow{x})-f(\overrightarrow{a}) \geqq -\sum_{i=1}^{m}\lambda_i g_i(\overrightarrow{x})+\sum_{i=1}^{m}\lambda_i g_i(\overrightarrow{a}) = -\sum_{i=1}^{m}\lambda_i g_i(\overrightarrow{x})$$

（$\lambda_i g_i(\overrightarrow{a})=0$も用いた）．

$\overrightarrow{x}$ は巻数達の定義域Sと不等式条件 $\overrightarrow{g} \leqq \overrightarrow{0}$ を満たす範囲で動かすから，特に$g_i(\overrightarrow{x}) \leqq 0$．KT条件から$\lambda_i \geqq 0$も再度用いると上の式の右辺は非負．よって$f(\overrightarrow{x})-f(\overrightarrow{a}) \geqq 0$, $\overrightarrow{x} \in S$, $\overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}$．

すなわち，f は不等式条件 $\overrightarrow{g} \leqq \overrightarrow{0}$ の下でSでの最小値を $\overrightarrow{a}$ で取る

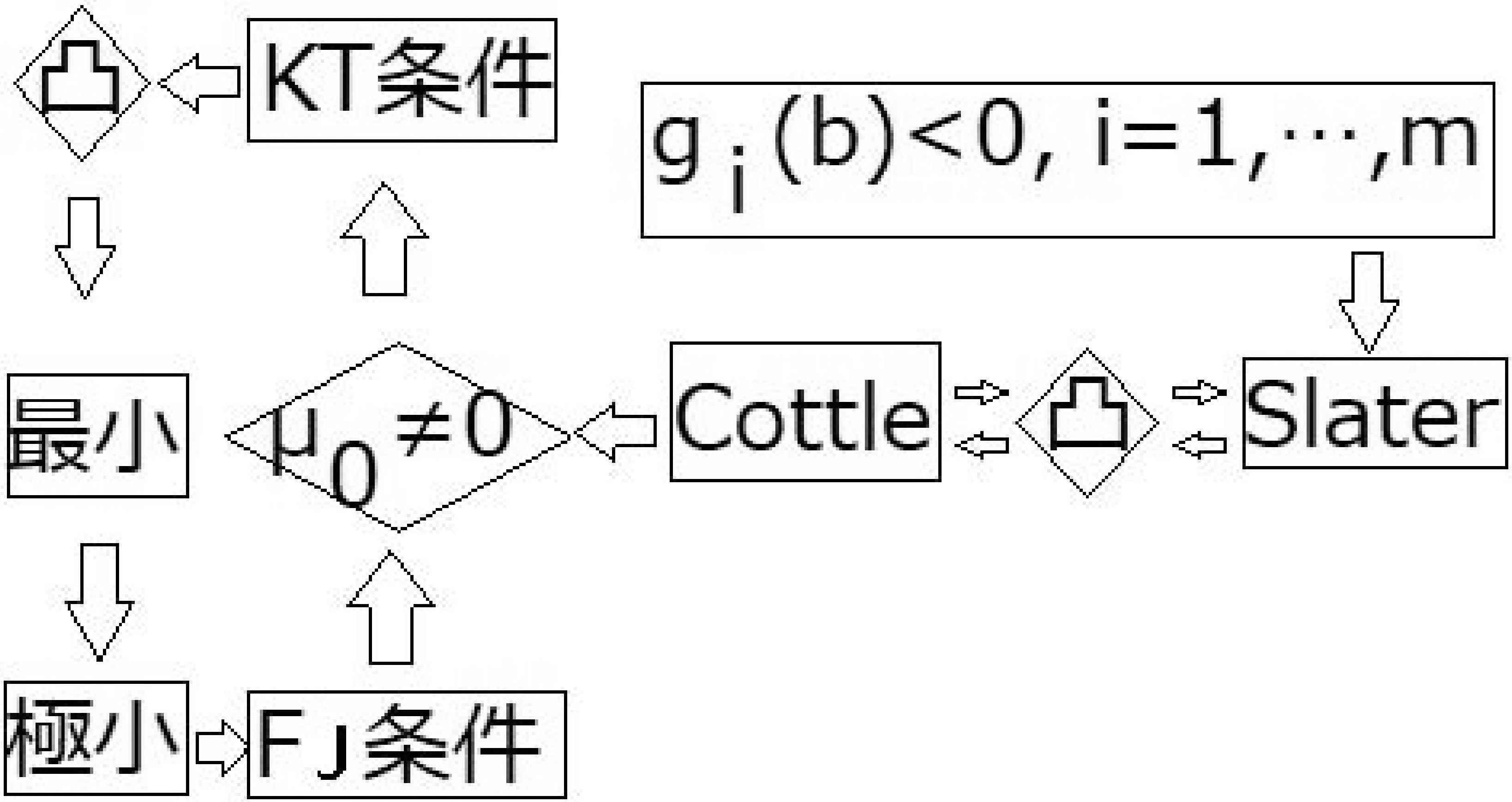
（証明終わり）．

不等式条件下の最小値とKT条件はいつ同値か

Cottleの制約想定は$I(\overrightarrow{a})$が$\overrightarrow{a}$ごとに違うので，最小値を取る点の候補ごとに成立しているか否か調べ直す必要があるが，「すべての$i = 1, \ldots, m$について$g_i(\overrightarrow{b}) < 0$が成り立つ」点$\overrightarrow{b} \in S$があれば，Cottleの制約想定はどこの点$\overrightarrow{a}$でも成り立つ．このことと先ほどの定理と秋の講義のFJ条件とKT条件と制約想定の関係から，

系．開凸集合$S \subset \mathbb{R}^n$を定義域とするm次元ベクトル値C^1級凸関数$\overrightarrow{g} : S \to \mathbb{R}^m$と$\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$を満たす点$\overrightarrow{a} \in S$について，$\overrightarrow{g}(\overrightarrow{b}) < \overrightarrow{0}$なる点$\overrightarrow{b} \in S$があるとき，
C^1級凸関数$f : S \to \mathbb{R}$が
$\overrightarrow{a}$で不等式条件$\overrightarrow{g} \leqq \overrightarrow{0}$の下での$S$での最小値を取ることと
$\overrightarrow{a}$でKuhn–Tucker条件が成り立つこと
は同値である．

不等式条件下の極値問題の諸条件



１０．分離定理（教科書p.48,49）

定理． nを自然数とする．n次元空間の集合$A \subset \mathbb{R}^n$と$B \subset \mathbb{R}^n$が，

AもBも凸集合で$A \cap B = \emptyset$を満たすならば，

$$\overrightarrow{x} \in A \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \geqq \alpha \text{ と } \overrightarrow{y} \in B \ \Rightarrow \ (\overrightarrow{y}, \overrightarrow{p}) \leqq \alpha$$

を満たす$\overrightarrow{p} \in \mathbb{R}^n$と$\alpha \in \mathbb{R}$がある．ここで$(\overrightarrow{u}, \overrightarrow{v})$は内積．

例１．離れている円板２つの間に直線を引くと円板は直線を挟んで違う半平面上
例２．離れた２球間に平面の仕切りを入れると球は仕切りを挟んで違う半空間内

n次元空間にも（超平面を適切に定義して）一般化できる

凸集合の定義（講義ｐｄｆ６の復習）

復習． n 次元空間内の集合（図形）$A \subset \mathbb{R}^n$ が凸集合とは，
A のどの２点のどの内分点も A の要素であること，言い換えると
A のどの２点を結ぶ線分も A に包含されること．
以上を式で書くと，A が凸集合とは
$(\forall \overrightarrow{x}, \overrightarrow{y} \in A)(\forall 0 < \lambda < 1)\ \lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y} \in A$，
すなわち，$(\forall \overrightarrow{x}, \overrightarrow{y} \in S)\ \{\lambda \overrightarrow{x} + (1-\lambda)\overrightarrow{y} \mid 0 \leqq \lambda \leqq 1\} \subset S$

忘れていたら，先に進む前に講義ｐｄｆ６の凸集合の部分を復習すること

餌を口に入れたパックマン

$n=2$（平面内の図形）の場合の**分離定理**：平面内の２つの凸集合 A と B に共有点がなければ $(A\cap B=\emptyset)$，A と B が直線を挟んで違う半平面にあるような直線が引ける．

例．離れている円板２つの間に直線を引くと円板は直線を挟んで違う半平面上
注目点．
1．A と B に共通点があれば分けられないので，$A\cap B=\emptyset$ は説明不要の必要条件．
2．餌 A を口に入れたパックマン B の図で A と B の間に直線を入れて分けられないので，凸集合の仮定はすなお．（イメージに不案内な場合は，たとえば pacman doodle で google 検索して開始を押して矢印キーで動かすと，暗殺教室のころせんせーが口を開けた所を横から見たような図形の画像が通路上の円を食べていくのを見ることができる.）
3．接していても共通接線で分けられることは（円の場合は高校の数学で）既知．
仮定 $A\cap B=\emptyset$ を維持するのがわかりやすいので，少なくとも一方の図形は接点を要素に持たない場合だけを分離定理では考える．
例．両方が閉集合ならば，接しない（距離が正の）場合だけを分離定理では考える．

超平面

重要． 平面だけでなく任意の n 次元ユークリッド空間 $\mathbb{R}^n$ の図形で成立！

分離定理． n 次元空間 $\mathbb{R}^n$ の中の共有点のない 2 つの凸集合は，**超平面**をうまく選ぶと，それによって分けられる**別々の半空間**に収められる．

凸集合は n 次元で定義済みだが，
平面のときの**分離するための直線と半平面は一般化**が必要．
n 次元空間の**超平面**：零ベクトルでない n 次元定ベクトル $\overrightarrow{p}$ と $\overrightarrow{x}_0$ があって，$(\overrightarrow{p},\ \overrightarrow{x} - \overrightarrow{x}_0) = 0$ を満たす点 $\overrightarrow{x}$ の集合．式で書くと
$H = \{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{p},\ \overrightarrow{x} - \overrightarrow{x}_0) = 0\}$ （ $(\cdot,\cdot)$ **は内積**）
例 $n = 2$．超平面は（前ページのとおり）直線．H は点 $\overrightarrow{x}_0$ を通り $\overrightarrow{p}$ に垂直な直線
例 $n = 3$．(高校数学で既習のとおり）超平面 H とは $\overrightarrow{x}_0$ を通り $\overrightarrow{p}$ を法線とする平面

半空間と分離定理

超平面 H が分離する半空間は（自然に符号で２分割して）

$\{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{p}, \overrightarrow{x} - \overrightarrow{x}_0) \geqq 0\}$ **と** $\{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{p}, \overrightarrow{x} - \overrightarrow{x}_0) \leqq 0\}$**．**
（前者は H から見て法線ベクトル $\overrightarrow{p}$ の向きの側の半空間，後者は反対側）
$\alpha = (\overrightarrow{p}, \overrightarrow{x}_0)$ **と置くと** $\{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{p}, \overrightarrow{x}) \geqq \alpha\}$ **と** $\{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{p}, \overrightarrow{x}) \leqq \alpha\}$**．**
この式を使うと前ページ最後の分離定理は今回の講義の冒頭のとおり：

分離定理． $A \subset \mathbb{R}^n$ **と** $B \subset \mathbb{R}^n$ **が凸集合で** $A \cap B = \emptyset$ **を満たすならば，**
$\overrightarrow{x} \in A \;\Rightarrow\; (\overrightarrow{x}, \overrightarrow{p}) \geqq \alpha$ **と** $\overrightarrow{x} \in B \;\Rightarrow\; (\overrightarrow{x}, \overrightarrow{p}) \leqq \alpha$ **を満たす**
$\overrightarrow{p} \in \mathbb{R}^n$ **と** $\alpha \in \mathbb{R}$ **がある．**

注１．集合では， $A \subset \{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{x}, \overrightarrow{p}) \geqq \alpha\}$ **と** $B \subset \{\overrightarrow{x} \in \mathbb{R}^n \mid (\overrightarrow{x}, \overrightarrow{p}) \leqq \alpha\}$**．**
注２． $\overrightarrow{p}$ **が法線ベクトルなら** $-\overrightarrow{p}$ **もなので，不等号** $\geqq \alpha$ **と** $\leqq \alpha$ **は入れ替え可能．**

２つの凸集合が接する場合について

・　凸集合 A と B は境界に条件はない（閉集合でも開集合でもどちらでもなくても分離定理は,「共有点のない凸集合」の条件だけで成立）.

・　特に，A と B が接しても，接点（双方にとって境界点）を少なくとも一方が要素に持たなければ $A \cap B = \emptyset$ が成立し分離定理成立.

例. $n = 2$, $A = \{(x, y) \mid (x+1)^2 + y^2 < 1\}$ と $B = \{(x, y) \mid (x-1)^2 + y^2 \leqq 1\}$ は,（直線 $x = 0$（y 軸）で分けられた左半平面 $U_1 = \{(x, y) \mid x \leqq 0\}$ と右半平面 $U_2 = \{(x, y) \mid x \geqq 0\}$ について $A \subset U_1$ と $B \subset U_2$ なので，前ページの分離定理の結論が $\overrightarrow{p} = (-1, 0)$ と $\alpha = 0$（$\overrightarrow{x}_0 = (0, 0)$）で成立）

・　接している場合は接線（接超平面）で仕切るので，証明が長くなる．Gordan の定理の証明で使うのはその場合だが，

・　閉集合どうしなら $A \cap B = \emptyset$ のとき（接点を共有できないので）離れている（接していない）から，分離定理の証明がやさしいので，なぜ成り立つかの理解のため，次ページ以降で閉集合の場合を証明.

接する場合の面倒な証明は教科書 p.47–50 を参照して，分離定理はありがたく使う.

原点と閉凸集合の分離定理の証明

定理０． n 次元空間の空で無い閉凸集合 $S \subset \mathbb{R}^n$ が $\overrightarrow{0} \notin S$ を満たすとき，$\overrightarrow{x} \in S \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \geqq d$ を満たす定ベクトル $\overrightarrow{p} \in \mathbb{R}^n$ と正定数 $d > 0$ の組がある．

注． 証明方法は他にもあるが，春に使い慣れた最大値の定理（春学期講義ｐｄｆ５参照）を用いる．そこで一時的に考察範囲を有界閉集合に限るために，最初にかってに１点 $\overrightarrow{c} \in S$ を固定する．

証明方針＝今の注の説明． 原点に一番近い S の点を $\overrightarrow{b}$ と置くと，$\overrightarrow{p} = \overrightarrow{b}$ と $d = (\overrightarrow{b}, \overrightarrow{p})$ が求める組．一番近い点は $f(\overrightarrow{x}) = \|\overrightarrow{x}\|^2$ の最小値を与える点で，最大値の原理が存在を保証する．

証明． $\overrightarrow{c} \in S$ を固定する．S と原点中心半径 $\|\overrightarrow{c}\|$ の球との共通部分 $R = S \cap \{\overrightarrow{x} \in \mathbb{R}^n \ | \ \|\overrightarrow{x}\| \ \leqq \ \|\overrightarrow{c}\|\}$ は有界閉集合なので最大値の原理から，$\overrightarrow{x} \ \in \ R$ に対して $\|\overrightarrow{x}\|^2 = f(\overrightarrow{x}) \geqq f(\overrightarrow{b}) = \|\overrightarrow{b}\|^2$ が成り立つ $\overrightarrow{b} \in R \subset S$ がある．特に $\overrightarrow{c}$ は球面上の S の点だから $\overrightarrow{c} \in R$ なので，$\|\overrightarrow{c}\|^2 \geqq \|\overrightarrow{b}\|^2$．球の外側の S の点 $\overrightarrow{x} \in S \cap R^c$ については，球の外側だからいまの不等式と合わせて $\|\overrightarrow{x}\|^2 \ > \ \|\overrightarrow{c}\|^2 \ \geqq \ \|\overrightarrow{b}\|^2$．よって球と関係なく，すべての $\overrightarrow{x} \in S$ に対して $\|\overrightarrow{x}\|^2 \geqq \|\overrightarrow{b}\|^2$．

S は凸集合なので（今回の講義１０冒頭の復習参照），$\overrightarrow{x} \in S$ かつ $0 < \lambda < 1$ ならば $\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{b} \in S$ だから，今の不等式の $\overrightarrow{x}$ をこれで置き換えることができる．このとき，左辺は自分自身との内積だから展開できて（以下各自確かめよ），

$$\lambda^2\|\overrightarrow{x}\|^2 + 2\lambda(1-\lambda)(\overrightarrow{x}, \overrightarrow{b}) + (1-\lambda)^2\|\overrightarrow{b}\|^2 \geqq \|\overrightarrow{b}\|^2.$$

移項して $\lambda > 0$ で割ると $\lambda\|\overrightarrow{x}\|^2 + 2(1-\lambda)(\overrightarrow{x}, \overrightarrow{b}) - (2-\lambda)\|\overrightarrow{b}\|^2 \geqq 0$. どんなに小さな正の λ でも成り立つには $(\overrightarrow{x}, \overrightarrow{b}) \geqq \|\overrightarrow{b}\|^2$ が必要．$\overrightarrow{p} = \overrightarrow{b}$ と $d = \|\overrightarrow{b}\|^2$ と選べば仮定 $\overrightarrow{0} \notin S$ と $\overrightarrow{b} \in R \subset S$ から $\overrightarrow{b} \neq \overrightarrow{0}$ なので $d > 0$（証明終わり）．

凸集合の差集合

分離定理の証明の次の段階で使う記号と命題を用意する．

n 次元空間 $\mathbb{R}^n$ の部分集合 $A \subset \mathbb{R}^n$ と $B \subset \mathbb{R}^n$ について差集合（それぞれの要素（ベクトル）の差をすべて集めた集合）を

$A - B = \{\overrightarrow{x} - \overrightarrow{y} \mid \overrightarrow{x} \in A,\ \overrightarrow{y} \in B\}$ と書く．

A と B が閉集合ならば $A - B$ も閉集合である（証明は解析学入門の範囲なので略）．

命題１．$A \subset \mathbb{R}^n$ と $B \subset \mathbb{R}^n$ が凸集合ならば，$A - B$ も凸集合．

命題２．$A \cap B = \emptyset$ ならば $A - B \not\ni \overrightarrow{0}$ （共有点のない集合の差集合の要素に原点（零ベクトル）は無い）．

命題１の証明．$A - B$ の任意の２点 $\overrightarrow{z}_i \in A - B$, $i = 1, 2$, は $\overrightarrow{x}_i \in A$ と $\overrightarrow{y}_i \in B$ を選んで $\overrightarrow{z}_i = \overrightarrow{x}_i - \overrightarrow{y}_i$, $i = 1, 2$, と書ける．これらと $0 < \lambda < 1$ に対して，

$$\lambda \overrightarrow{z}_1 + (1-\lambda)\overrightarrow{z}_2 = (\lambda \overrightarrow{x}_1 + (1-\lambda)\overrightarrow{x}_2) - (\lambda \overrightarrow{y}_1 + (1-\lambda)\overrightarrow{y}_2)$$

と書けて，A と B は仮定により凸集合だから $\lambda \overrightarrow{x}_1 + (1-\lambda)\overrightarrow{x}_2 \in A$ と $\lambda \overrightarrow{y}_1 + (1-\lambda)\overrightarrow{y}_2 \in B$ が成り立つので，$\lambda \overrightarrow{z}_1 + (1-\lambda)\overrightarrow{z}_2 \in A - B$．よって，凸集合の定義（今回の講義冒頭の復習）から $A - B$ は凸集合である（証明終わり），

命題２の証明．背理法で，$\overrightarrow{0} \in A - B$ とすると $\overrightarrow{x} = \overrightarrow{y}$ となる $\overrightarrow{x} \in A$ と $\overrightarrow{y} \in B$ があるが，$\overrightarrow{x} = \overrightarrow{y} \in A \cap B$ を意味するので $A \cap B = \emptyset$ に矛盾する（証明終わり）．

閉凸集合どうしの場合の分離定理の証明

閉凸どうしの分離定理． $A \subset \mathbb{R}^n$ も $B \subset \mathbb{R}^n$ も閉凸集合で $A \cap B = \emptyset$ ならば，
$\overrightarrow{x} \in A \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \geqq d + \alpha$ と $\overrightarrow{x} \in B \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \leqq \alpha$ を満たす
$\overrightarrow{p} \in \mathbb{R}^n$ と $d > 0$ と $\alpha \in \mathbb{R}$ がある．
注． 後で使うのは接する（一方が閉でない）場合なので（講義冒頭の定理のとおり）
$d = 0$ だが，閉凸のときは A と B が離れているので強く，$d > 0$ なので，残しておく．
証明． A と B が凸なので前ページ命題 1 から $A - B$ も凸で，命題 2 から $\overrightarrow{0} \notin A - B$．
よって前々ページ定理 0 で $S = A - B$ と置くと仮定が満たされるので，結論から
「$\overrightarrow{z} \in A - B \ \Rightarrow \ (\overrightarrow{z}, \overrightarrow{p}) \geqq d$」を満たす定ベクトル $\overrightarrow{p} \in \mathbb{R}^n$ と正定数 $d > 0$ の
組がある．よって $\overrightarrow{x} \in A$ かつ $\overrightarrow{y} \in B$ ならば

$$(\overrightarrow{x}, \overrightarrow{p}) \geqq d + (\overrightarrow{y}, \overrightarrow{p}), \quad \overrightarrow{x} \in A, \ \overrightarrow{y} \in B \qquad (*)$$

が成り立つ．$\overrightarrow{x}$ を固定すると (*) から $\{(\overrightarrow{y}, \overrightarrow{p}) \mid \overrightarrow{y} \in B\} \subset \mathbb{R}$ は上に有界なので
上限 $\alpha = \sup\{(\overrightarrow{y}, \overrightarrow{p}) \mid \overrightarrow{y} \in B\} \in \mathbb{R}$ があり，上限の定義（解析学入門参照）と
(*) から $(\overrightarrow{x}, \overrightarrow{p}) \geqq d + \alpha$ と $\alpha \geqq (\overrightarrow{y}, \overrightarrow{p})$ が任意の $\overrightarrow{x} \in A$ と $\overrightarrow{y} \in B$ に対して成り立つ（証明終わり）．
注． 閉集合と仮定したので $\alpha = (\overrightarrow{y}_0, \overrightarrow{p})$ を満たす $\overrightarrow{y}_0 \in B$ がある（最大値の原理と類似の性質）．$\overrightarrow{x}_0 \in A$ を選んで $\alpha = (\overrightarrow{x}_0, \overrightarrow{p}) - d$ とおくことも可能．一般の場合で接するとき（$d = 0$）は $\overrightarrow{x}_0$ と $\overrightarrow{y}_0$ は接点．

１１．Gordanの定理

定理（講義ｐｄｆ１既出）．$\overrightarrow{a}_1, \ldots, \overrightarrow{a}_m \in \mathbb{R}^n$ とするとき，(1) と (2) は同値：

(1)　非負要素たちからなる $\overrightarrow{0}$ でない m 次元ベクトル $\overrightarrow{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix}$

であって $p_1\overrightarrow{a}_1 + \cdots + p_m\overrightarrow{a}_m = \overrightarrow{0}$ を満たすものがある．

(2)　内積 $(\overrightarrow{a}_i, \overrightarrow{y}) < 0,\ i = 1, \ldots, m,$ を満たす $\overrightarrow{y} \in \mathbb{R}^n$ はない．

- 証明は (1) → (2) は代入，(2) → (1) は分離定理【比喩：(2) が方程式，(1) が解】
- $A = (\overrightarrow{a}_1\ \cdots\ \overrightarrow{a}_m)$ と置いて $V = \{{}^tA\overrightarrow{y} \mid \overrightarrow{y} \in \mathbb{R}^n\}$
- A による写像の値域 V に分離定理【内積の正負を見るための抽象】
- 分離定理は不等式だが線形空間 V に使うと等式

数学的に良い形への書き換え

行列表記のGordanの定理． A が $n\times m$ 行列のとき(1)(2)が同値：
(1) $A\overrightarrow{p}=\overrightarrow{0}$, $\overrightarrow{p}\geqq\overrightarrow{0}$, $\overrightarrow{p}\neq\overrightarrow{0}$ を満たす $\overrightarrow{p}\in\mathbb{R}^m$ がある．
(2) ${}^tA\overrightarrow{y}\in\mathbb{R}^m$ が全成分負になる $\overrightarrow{y}\in\mathbb{R}^n$ がない．${}^tA$ は転置行列．

書き換えに過ぎないことの証明． 行列 A が与えられたとき「列の小箱」にわけて $A=(\overrightarrow{a}_1\ \cdots\ \overrightarrow{a}_m)$ と置くと $A\overrightarrow{p}=p_1\overrightarrow{a}_1+\cdots+p_m\overrightarrow{a}_m$ となるからこのページの(1)から前ページの(1)を得る．また，列ベクトルの内積を行列の積で書くと $(\overrightarrow{a},\overrightarrow{y})={}^t\overrightarrow{a}_i\overrightarrow{y}$ だから ${}^tA\overrightarrow{y}=\begin{pmatrix}(\overrightarrow{a}_1,\overrightarrow{y})\\ \vdots\\ (\overrightarrow{a}_m,\overrightarrow{y})\end{pmatrix}$ なので，このページの(2)から前ページの(2)を得る．逆に前ページのように列ベクトル $\overrightarrow{a}_1,\ldots,\overrightarrow{a}_m\in\mathbb{R}^n$ が与えられたとき，これらを列とする行列を $A=(\overrightarrow{a}_1\ \cdots\ \overrightarrow{a}_m)$ と置くと，同様に前ページの(1)(2)からそれぞれこのページの(1)(2)を得る（証明終わり）．

Gordanの定理の注．

- 小学校で鶴亀算，中学校で連立１次方程式と呼ぶ対象を大学で $A\overrightarrow{x}=\overrightarrow{b}$ と書いて $\overrightarrow{x}=A^{-1}\overrightarrow{b}$ を得るように，Gordanの定理も前ページの行列表記が見通し良い．
- 前ページの形は，前々ページの（中学の連立方程式と大学の行列表記と同程度の）単なる書き換えであることが，前ページの証明で分かった．
- よって，前ページと前々ページの定理を区別しない（断らずに前ページの形で定理を証明し，FJ条件の導出では前々ページの形を使う）．

- 定理の主張は (1)⇔(2)．この同値の気持ちは，方程式と解の同値($A\overrightarrow{x}=\overrightarrow{b}\Leftrightarrow\overrightarrow{x}=A^{-1}\overrightarrow{b}$) と類似のレベル：
- (1)⇒(2) は証明容易（解を方程式に代入して成立を確かめるレベル）
- (2)⇒(1) からFJ条件を得る（方程式から解を得るレベル．技が必要）
- 技＝分離定理．(1) の $\overrightarrow{p}$ が，方程式から見いだす解に対応．

－ 見え見えのこと：分離定理の $\overrightarrow{p}$ がGordanの定理の $\overrightarrow{p}$

－ 注目点１：分離定理の２つの凸集合の選択が，昔の偉い人たちのアイデア

－ 注目点２：分離定理は不等号しか結論しないが，Gordanの定理 (1) は等号 !?

○ 一方の凸集合が線形空間（凸かつ凹）だから等号を得る（以下で証明）

Gordanの(1)⇒(2)の証明

注. やさしいけど使わない（解を方程式に代入して成り立つと言うような内容）.
・　解を方程式に代入して確かめるのと類似の価値はあるので証明を確認しておく.

(1)⇒(2)の証明.
(1)の$A\overrightarrow{p} = \overrightarrow{0}$とどんなベクトル$\overrightarrow{y} \in \mathbb{R}^n$と内積を取っても0. 行列の積の表記で${}^t\overrightarrow{y}A\overrightarrow{p} = 0$.
内積は数だから転置を取っても同じ：${}^t\overrightarrow{p}({}^tA\overrightarrow{y}) = 0$.
成分で書くと$p_1({}^tA\overrightarrow{y})_1 + \cdots + p_m({}^tA\overrightarrow{y})_m = 0$.
どのiも$p_i \geqq 0$で1つ以上正という(1)の仮定なので，背理法で，
$({}^tA\overrightarrow{y})_i$たちが全部負とすると，左辺は負になって上記「$= 0$」と矛盾.
よって(2)のとおり，「${}^tA\overrightarrow{y}$が全成分負」ということは
どんな$\overrightarrow{y} \in \mathbb{R}^n$を持ってきても起きない.（証明終わり）

分離定理　から　Gordanの(2)⇒(1)

Gordanの定理の(2)⇒(1)の証明.
分離定理（講義ｐｄｆ１０）のnをm，AをV，BをTと書き直して
再掲：『$V \subset \mathbb{R}^m$も$T \subset \mathbb{R}^m$も凸集合で$V \cap T = \emptyset$を満たすならば，
$\overrightarrow{x} \in V \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \geqq \alpha$ と $\overrightarrow{x} \in T \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \leqq \alpha$
を満たす$\overrightarrow{p} \in \mathbb{R}^m$と$\alpha \in \mathbb{R}$がある.』

Gordanの定理の行列表記の$A = (\overrightarrow{a}_1 \ \ldots \ \overrightarrow{a}_m)$を用いて，
$V = \{{}^tA\overrightarrow{y} \mid \overrightarrow{y} \in \mathbb{R}^n\}$と$T = \{\overrightarrow{x} \in \mathbb{R}^m \mid$ 全座標負 $\}$ に上記分離定理を使う.
例. $m = 2$（平面）のときTは第３象限（軸を含まない）.

・　Vは（和とスカラー倍について閉じているので）線形空間. 線形空間と第３象限は凸集合（講義ｐｄｆ６）.
・　(2)の仮定『${}^tA\,\overrightarrow{y} \in \mathbb{R}^m$が全成分負になる$\overrightarrow{y} \in \mathbb{R}^n$がない』から$V \cap T = \emptyset$.
よって分離定理の仮定が成り立つからその結論が成り立つ（上記「再掲」参照）.

（続）Gordanの定理の(2)⇒(1)の証明

前ページの結論（再掲）：Gordanの(2)が成り立つ時$V = \{{}^tA\overrightarrow{y} \mid \overrightarrow{y} \in \mathbb{R}^n\}$と
$T = \{\overrightarrow{x} \in \mathbb{R}^m \mid$ 全座標負$\}$について，$\overrightarrow{p} \in \mathbb{R}^m$があって，
$\overrightarrow{x} \in V \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \geqq \alpha$ と $\overrightarrow{x} \in T \ \Rightarrow \ (\overrightarrow{x}, \overrightarrow{p}) \leqq \alpha$ が成り立つ．

- ここから：

線形空間と「第３象限」なので分離定理の一般論よりも強い結論．

– $\overrightarrow{0} = {}^tA\overrightarrow{0} \in V$だから上記「$\overrightarrow{x} \in V \Rightarrow \cdots$」を$\overrightarrow{x} = \overrightarrow{0}$に適用できて$0 = (\overrightarrow{0}, \overrightarrow{p}) \geqq \alpha$．

– T（平面$m = 2$なら第３象限）は全成分負なすべての点を集めたので，原点$\overrightarrow{0}$のいくらでも近い点について$(\overrightarrow{x}, \overrightarrow{p}) \leqq \alpha$ よって$\alpha \geqq 0$でないといけない．

- 以上から，$\alpha = 0$と決まる．つまり，

$\overrightarrow{x} \in V \Rightarrow {}^t\overrightarrow{x}\,\overrightarrow{p} = (\overrightarrow{x}, \overrightarrow{p}) \geqq 0$．

（左辺は行列の式で中央の辺の内積を書き直したもの．）

Gordanの定理の(2)⇒(1)の証明（終わり）

前ページの結論（再掲）：Gordanの(2)が成り立つ時 $V = \{{}^t A \overrightarrow{y} \mid \overrightarrow{y} \in \mathbb{R}^n\}$ について $\overrightarrow{p} \in \mathbb{R}^m$ があって，$\overrightarrow{x} \in V \Rightarrow {}^t\overrightarrow{x}\,\overrightarrow{p} = (\overrightarrow{x}, \overrightarrow{p}) \geqq 0$．

- V の定義（1行目の**再掲**）から $\overrightarrow{x} \in V$ とは $\overrightarrow{x} = {}^t A \overrightarrow{y}$ の形．
- 合わせると，$(\overrightarrow{y}, A\overrightarrow{p}) = {}^t\overrightarrow{y} A \overrightarrow{p} = {}^t\overrightarrow{x}\,\overrightarrow{p} \geqq 0$．
- $\overrightarrow{y} \in \mathbb{R}^n$ は何でも良いので $\overrightarrow{y} = -A\overrightarrow{p}$ と選ぶと

$-\|A\overrightarrow{p}\|^2 = (-A\overrightarrow{p}, A\overrightarrow{p}) = (\overrightarrow{y}, A\overrightarrow{p}) \geqq 0$

ノルムは非負なのでこれが成り立つのは等号の場合だけ．よって

$\|A\overrightarrow{p}\| = 0$．

ノルム（長さ）が0のベクトルは零ベクトルだけだから，$A\overrightarrow{p} = \overrightarrow{0}$．

よって(2)から(1)を得た（証明終わり）．

１２．GordanからFJへ（教科書p.27,28）

★今回は，新しい問はありません．前回の略解のみです．

m と n は正整数．n 変数 C^1 級実数値関数たち f，g_1，g_2，$\dots$，g_m

$\overrightarrow{a} \in \mathbb{R}^n$ は $g_i(\overrightarrow{a}) \leqq 0$, $i = 1, \dots, m$, を満たす点．

以上を今回の講義では１組（与えられたとして）以下固定する．

Fritz–John条件（講義ｐｄｆ１既出）：　$\exists \mu_0, \mu_1, \dots, \mu_m$;

$\mu_0 \nabla f(\overrightarrow{a}) + \mu_1 \nabla g_1(\overrightarrow{a}) + \cdots + \mu_m \nabla g_m(\overrightarrow{a}) = \overrightarrow{0}$,

$\mu_1 g_1(\overrightarrow{a}) = \cdots = \mu_m g_m(\overrightarrow{a}) = 0$,

$\mu_0 \geqq 0$, $\mu_1 \geqq 0$, $\dots$, $\mu_m \geqq 0$, $(\mu_0, \mu_1, \cdots, \mu_m) \neq \overrightarrow{0}$.

不等式条件下の極小値と Fritz–John 条件

定理２（講義ｐｄｆ１既出）． 前ページの関数たちについて，条件 $\overrightarrow{g} \leqq 0$ の下での極小値を f が $\overrightarrow{a}$ で取るならば，前ページの $\overrightarrow{a}$ での Fritz-John 条件（以下講義内略記で FJ 条件）が成り立つ．

Gordan の定理の (2) ⇒ (1) と次の補題を用いることでこの定理が証明できる． 前ページのとおり引き続き f や $\overrightarrow{a}$ は前ページの青字を満たす．

下り坂の補題． 0 での微分 $\overrightarrow{x}'(0)$ が存在する関数 $\overrightarrow{x}: \mathbb{R} \to \mathbb{R}^n$ が $\nabla f(\overrightarrow{x}(0))\overrightarrow{x}'(0) < 0$（勾配ベクトルと微分の内積の行列表記）を満たせば，
$\exists t_0 > 0;\ (\forall 0 < t < t_0)\ f(\overrightarrow{x}(t)) < f(\overrightarrow{x}(0))$．
特に，$\overrightarrow{u} \in \mathbb{R}^n$ が $\nabla f(\overrightarrow{a})\overrightarrow{u} < 0$ （これも内積の積表記）を満たせば，
$\exists t_0 > 0;\ (\forall 0 < t < t_0)\ f(\overrightarrow{a} + t\overrightarrow{u}) < f(\overrightarrow{a})$．

注１，後半は前半で $\overrightarrow{x}(t) = \overrightarrow{a} + t\overrightarrow{u}$ と選んだ場合なので，前半を証明すれば良い．
注２，$\nabla f(\overrightarrow{a})$ は地点 $\overrightarrow{a}$ での最も急な登りの方向（講義ｐｄｆ２，講義ｐｄｆ１の問）なので，補題の意味は，それと鈍角の方向に進めば f の値はしばらく小さいということ（この講義の中だけの名前で下り坂の補題）．以下，そのことを数式 (ϵ–δ) で証明する．数学的な注意点は，でこぼこした道を下山する可能性があるが，関係なく $\overrightarrow{a}$ の十分近くでは $f(\overrightarrow{a})$ より低いことの証明（平均値の定理経由ではなく，微分の定義から直接証明，という意味）

補題の証明

下り坂の補題の証明． $h = f \circ \overrightarrow{x}:\ \mathbb{R} \to \mathbb{R}$と置くと，合成関数の微分法則（教科書第２章定理4(1)，春学期講義ｐｄｆ７）から
$c = h'(0) = \nabla f(\overrightarrow{x}(0))\overrightarrow{x}'(0)$
が存在$(c \in \mathbb{R})$．仮定から$c < 0$．tの１変数関数hの微分の定義（微分積分，解析学入門）を分母を払って書き下すと，
$(\forall \epsilon > 0)\exists \delta > 0;\ (\forall |t| < \delta)\ |h(t) - h(0) - c\,t| < \epsilon\,t$．
ϵは任意の正数でよいので$\epsilon = -\frac{1}{2}c > 0$と選び，その選択に対応する$\delta$を$t_0 = \delta > 0$と置いて，絶対値を外して，さらに片側の不等号だけ残すと（各自確かめよ），
$(\forall |t| < t_0)\ h(t) < h(0) + c\,t - \frac{1}{2}c\,t = \frac{1}{2}c\,t$．
特に$c < 0$なので$0 < t < t_0$のとき$f(\overrightarrow{x}(t)) = h(t) < h(0) = f(\overrightarrow{x}(0))$が成り立つ（証明終わり）．

Gordanの定理の(2)⇒(1)

Gordanの定理(2)⇒(1)（講義ｐｄｆ１１参照）．

$\overrightarrow{a}_1, \ldots, \overrightarrow{a}_m \in \mathbb{R}^n$について，

内積を$(\cdot,\cdot)$と書くとき，$(\overrightarrow{a}_i, \overrightarrow{y}) < 0,\ i = 1, \ldots, m,$ を満たす $\overrightarrow{y} \in \mathbb{R}^n$が無いことがわかれば，

$$\overrightarrow{p} \geqq \overrightarrow{0} \text{ かつ } \overrightarrow{p} \neq \overrightarrow{0} \text{ を満たす } \overrightarrow{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix} \in \mathbb{R}^m$$

（非負要素からなる零ベクトルでないm次元ベクトル）であって，

$p_1\overrightarrow{a}_1 + \cdots + p_m\overrightarrow{a}_m = \overrightarrow{0}$ を満たすものがある．

不等式条件下の極小値問題でのＦＪ条件の証明

３ページ前の定理２の証明．

- $I(\overrightarrow{a}) = \{i \in \{1, \ldots, m\} \mid g_i(\overrightarrow{a}) = 0\}$ **と置く．**

意味． $\overrightarrow{a}$ で不等式条件のうち等号 $(g_i(\overrightarrow{a}) = 0)$ が成り立つ添字 i の集合．

最初のページの青字設定で $\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$ だから $i \notin I(\overrightarrow{a}) \Leftrightarrow g_i(\overrightarrow{a}) < 0$．

g_i は C^1 なので C^0（連続）だから，$g_i(\overrightarrow{a}) < 0$ のとき $\overrightarrow{a}$ のある近傍で $g_i(\overrightarrow{x}) < 0$．共通近傍の中の点 $\overrightarrow{x}$ で $g_i(\overrightarrow{x}) < 0$, $i \notin I(\overrightarrow{a})$．

$\overrightarrow{a}$ の近傍で成り立つので，$\overrightarrow{a}$ で極小を取るかという問題はこれらの条件が無いときと同一の問題．

以下，**不等式条件は** $g_i \leqq 0$, $i \in I(\overrightarrow{a})$, **だけ**を考えれば良い．

（続）ＦＪ条件の証明（補題の適用）

- $\nabla f(\overrightarrow{a})\overrightarrow{u} < 0,\ \nabla g_i(\overrightarrow{a})\overrightarrow{u} < 0,\ i \in I(\overrightarrow{a}),$ を同時に満たす $\overrightarrow{u}$ が無いことを今から背理法で証明する（あるとして矛盾を導く）．

下り坂の補題を各々の不等式に適用して合わせることで（t_0 をこれらの中で最小のものに選ぶことで），$t_0 > 0$ が存在して，$0 < t < t_0$ ならば $f(\overrightarrow{a}+t\overrightarrow{u}) < f(\overrightarrow{a}),\ g_i(\overrightarrow{a}+t\overrightarrow{u}) < g_i(\overrightarrow{a}) = 0,\ i \in I(\overrightarrow{a}),$ がすべて成り立つ．

すなわち，$0 < t < t_0$ ならば点 $\overrightarrow{a} + t\overrightarrow{u}$ は不等式条件をすべて満たし，そこで f の値は $f(\overrightarrow{a})$ より小さいので，f は $\overrightarrow{a}$ で極小値ではない．これは定理の仮定（極小値を取る）に矛盾する．背理法によって，先の不等式達が同時に成り立つ $\overrightarrow{u}$ は存在しない．

（続）FJ条件の証明（Gordanの定理の適用）

$I(\overrightarrow{a})$ の要素の数に1加えた数を m と置いた上で，一時的に関数たち f，g_i，$i \in I(\overrightarrow{a})$，を通し番号で h_1，$\ldots$，h_m と置き，
$\overrightarrow{a}_i = {}^t\nabla h_i(\overrightarrow{a})$，$i = 1, 2, \ldots, m$，と置く．
注．転置は，教科書の勾配ベクトルは行ベクトル，位置ベクトル等は列ベクトルだから．元の極値問題とGordanの定理で別の値を同じ m と置いたのは教科書の責任．

以上の記号に加えて前ページの結論で $\overrightarrow{y} = \overrightarrow{u}$ とすると，3ページ前のGordanの定理(2)⇒(1)の仮定が成り立つ．その結論から，$\overrightarrow{p} \geqq \overrightarrow{0}$ かつ $\overrightarrow{p} \neq \overrightarrow{0}$ を満たす $\overrightarrow{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix} \in \mathbb{R}^m$ であって，
$p_1\overrightarrow{a}_1 + \cdots + p_m\overrightarrow{a}_m = \overrightarrow{0}$ を満たすものがある．
最後の式は，ここでの通し番号の関数の値に戻して転置を外すと
（零ベクトルは縦も横も転置記号を付けないことにして）

$$p_1\nabla h_1(\overrightarrow{a}) + p_2\nabla h_2(\overrightarrow{a}) + \cdots + p_m\nabla h_m(\overrightarrow{a}) = \overrightarrow{0}$$

FJ条件の証明（終わり）

前ページ最後の行の h_i たちを本来の f と g_i たちに戻すと通し番号の p_i たちと添字がずれる（g_i は等号 $g_i(\overrightarrow{a}) = 0$ を満たすものだけ並べていたので添字が飛び飛び）．$h_i = g_j$ のとき，$p_i h_i = \mu_j g_j$ と添字を揃えるために $\mu_j = p_i$ と添字を付け直した係数を μ_j たちとし，$\mu_0 = p_1$ として添字が衝突しないようにすると，前ページ最後の行は $\mu_0 \nabla f(\overrightarrow{a}) + \sum_{i \in I(\overrightarrow{a})} \mu_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0}$．

$i \notin I(\overrightarrow{a})$ なる μ_i はこの式にない（定義されていない）ので $\mu_i = 0,\ i \notin I(\overrightarrow{a})$，と 0 に置いてもかまわない（しかも便利）．これでFJ条件のベクトル等式

$$\mu_0 \nabla f(\overrightarrow{a}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\overrightarrow{a}) = \overrightarrow{0}$$ を得る．

$I(\overrightarrow{a})$ の定義から $i \in I(\overrightarrow{a}) \Leftrightarrow g_i(\overrightarrow{a}) = 0$ で，今の μ_i の選択から $i \notin I(\overrightarrow{a}) \Rightarrow \mu_i = 0$ だから，
$\mu_1 g_1(\overrightarrow{a}) = \cdots = \mu_m g_m(\overrightarrow{a}) = 0$．
前ページの $p_i \geqq 0$ たちから添字を付け直して $i \in I(\overrightarrow{a}) \cup \{0\} \Rightarrow \mu_i \geqq 0$ であり，$i \notin I(\overrightarrow{a}) \Rightarrow \mu_i = 0$ と置いたから，$\mu_i \geqq 0,\ i = 0, 1, \ldots, m$．同じく前ページから p_i の中に 0 でないものがあるから，μ_i **たちの中に 0 でない**ものがある．
よってFJ条件を得た（証明終わり）．

１３．分離定理：経済学への応用例

◯関数の値域の分離定理

－　経済学では関数の最大最小に興味がある

・鞍点定理（今回の講義）

・二人零和非協力ゲームのナッシュ均衡の存在定理（確率論入門２）

◯線形空間と「第３象限」の分離定理

－　分離定理は不等号 $(\overrightarrow{p} \geqq \overrightarrow{0})$ だが等号を得る原理があると等号

・不等式条件下の最小値問題（本講義経済数学２既習）

・数理ファイナンスの第１基本定理（ファイナンス数学）

◯個別には不等号でも集合としての等号は成立

・ベイズ解集合と許容解集合の一致定理（確率論入門２）

鞍点定理（教科書 p.52）

教科書に従って，制約想定下の m 個の n 変数凸関数不等式条件下の凸関数の最小値を取る点についての鞍点定理の形で定理を書く． 凸集合 $S \subset \mathbb{R}^n$，
凸関数たち $f:\ S \to \mathbb{R}$ と $\overrightarrow{g}:\ S \to \mathbb{R}^m$ （$\overrightarrow{g}$ は各成分毎に凸）
$g_i(\overrightarrow{x}_0) < 0,\ i = 1, \ldots, m,$ となる $\overrightarrow{x}_0 \in S$ があるとする（Slater の制約想定の十分条件，つまり KT 条件の解が唯一存在して f がそこで条件下最小値）

$L:\ S \times \mathbb{R}^m \to \mathbb{R};\ L(\overrightarrow{x}, \overrightarrow{y}) = f(\overrightarrow{x}) + (\overrightarrow{y}, \overrightarrow{g}(\overrightarrow{x}))$
（右辺第２項の $(\cdot, \cdot)$ は内積．ルジャンドル変換の母関数， ラグランジュ関数）

鞍点定理． $\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$ を満たす $\overrightarrow{a} \in S$ について (i) と (ii) は同値：

(i) $\displaystyle\min_{\overrightarrow{x} \in S;\ \overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}} f(\overrightarrow{x}) = f(\overrightarrow{a})$ （つまり $f(\overrightarrow{a})$ が条件下の最小値），

(ii) $\exists \overrightarrow{b} \in \mathbb{R}^m_+;\ L(\overrightarrow{x}, \overrightarrow{b}) \geqq L(\overrightarrow{a}, \overrightarrow{b}) \geqq L(\overrightarrow{a}, \overrightarrow{y}),\ \overrightarrow{x} \in S,$
$\overrightarrow{y} \geqq \overrightarrow{0}$ （鞍点）．

さらに，いずれか（したがって両方）が成り立つとき
(iii) $b_i\, g_i(\overrightarrow{a}) = 0,\ i = 1, 2, \ldots, m$． 特に $L(\overrightarrow{a}, \overrightarrow{b}) = f(\overrightarrow{a})$．

関数の値域の分離定理

分離定理：「膨らんだ風船２個の間に板の仕切りを入れられる」

（餌を食べるパックマンは直線で餌を仕切れない）

理論への応用では関数の値域の部分集合に分離定理を使う．（財の関数の効用や利潤などを最大化つまり値の研究が元々の理論の目的だから当然）

$S \subset \mathbb{R}^n$ 上の k 成分関数 $\overrightarrow{h} : \ S \to \mathbb{R}^k$ の**像**

$\mathrm{Im}\,\overrightarrow{h} := \{\overrightarrow{h}(\overrightarrow{x}) \mid \overrightarrow{x} \in S\} \subset \mathbb{R}^k$ （$\overrightarrow{h}$ **が取る値**の組を集めた集合）

値域での分離定理． S が凸，$\overrightarrow{h}$ が凸，かつ $\displaystyle\min_{\overrightarrow{x} \in S} \max_{i=1,\dots,k} h_i(\overrightarrow{x}) \geqq 0$

のとき，$\exists \overrightarrow{p} \in \mathbb{R}^k_+ \setminus \{\overrightarrow{0}\}; \ \displaystyle\min_{\overrightarrow{x} \in S} (\overrightarrow{p}, \overrightarrow{h}(\overrightarrow{x})) \geqq 0$ が成り立つ．

注：結論の $(\cdot, \cdot)$ は内積，

値域の分離定理と既出の分離定理の関係

・値域の分離定理の仮定 $\displaystyle\min_{\overrightarrow{x}\in S}\max_{i=1,\dots,k} h_i(\overrightarrow{x}) \geqq 0$ は定義域 S のどこでも「$\overrightarrow{h}$ の値（ベクトル）の全成分が負」は起きない，ということだから値域空間 $\mathbb{R}^k$ の全成分負のベクトルの集合 T （$k=2$ なら第３象限）と像 $\mathrm{Im}\,\overrightarrow{h}$ が共通部分を持たないということ．

・結論 $\displaystyle\min_{\overrightarrow{x}\in S}(\overrightarrow{p}, \overrightarrow{h}(\overrightarrow{x})) \geqq 0$ は非負成分の $\overrightarrow{p} \neq \overrightarrow{0}$ があって，S のどこでも $\overrightarrow{h}$ の値（ベクトル）と内積が非負ということ．

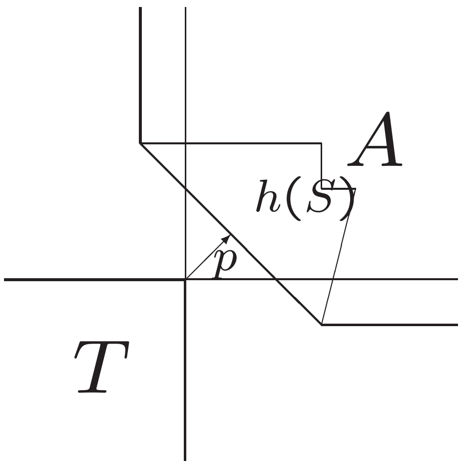
→　**講義 11 の Gordan の定理の (2) ⇒ (1) の証明と形は合っている！**

☆ただし，像 $\mathrm{Im}\,\overrightarrow{h}$ が凸集合とは限らない．分離定理の集合 A は $\mathrm{Im}\,\overrightarrow{h}$ を包含する凸集合に選ぶ必要（次ページの証明冒頭の「$\geqq$」参照）

・Gordan の定理の証明の $A = V = \{{}^t A\overrightarrow{y} \mid \overrightarrow{y} \in \mathbb{R}^n\}$ は，$S = \mathbb{R}^n$ と $\overrightarrow{h}(\overrightarrow{x}) = {}^t A\overrightarrow{x}$ と置いた場合の値域の分離定理になる．(講義 11 では V を天下りで上記のように置いたが，その種明かし．)

関数の値域での分離定理の証明

$A = \{\overrightarrow{u} \in \mathbb{R}^k \mid \exists \overrightarrow{x} \in S;\ \ \overrightarrow{u} \geqq \overrightarrow{h}(\overrightarrow{x})\}$ および $T = \{\overrightarrow{u} \in \mathbb{R}^k \mid u_i < 0,\ i = 1, \ldots, k\}$ とおくと，前ページの注のとおり $A \cap T = \emptyset$．



T は凸（講義pdf６例4）で，A も凸である．なぜなら $\overrightarrow{u}, \overrightarrow{v} \in A$ と $0 < \lambda < 1$ に対して，$\overrightarrow{u} \geqq \overrightarrow{h}(\overrightarrow{x})$ および $\overrightarrow{v} \geqq \overrightarrow{h}(\overrightarrow{y})$ が成り立つ $\overrightarrow{x}, \overrightarrow{y} \in S$ が存在し，S が凸だから $\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y} \in S$ であって，$\overrightarrow{h}: S \to \mathbb{R}^k$ が凸だから，

$$\begin{aligned}\lambda\overrightarrow{u} + (1-\lambda)\overrightarrow{v} &\geqq \lambda\overrightarrow{h}(\overrightarrow{x}) + (1-\lambda)\overrightarrow{h}(\overrightarrow{y}) \\ &\geqq \overrightarrow{h}(\lambda\overrightarrow{x} + (1-\lambda)\overrightarrow{y})\end{aligned}$$

となって，$\lambda\overrightarrow{u} + (1-\lambda)\overrightarrow{v} \in A$ となるからである．

よって分離定理（講義pdf１０）から，$(\overrightarrow{p}, \overrightarrow{u}) \geqq c \geqq (\overrightarrow{p}, \overrightarrow{v}),\ \ \overrightarrow{u} \in A,\ \ \overrightarrow{v} \in T.$ となる $\overrightarrow{p} \in \mathbb{R}^k \setminus \{\overrightarrow{0}\}$ と $c \in \mathbb{R}$ の組がある．

T の中で $v_i \to -\infty$ とできるので，$p_i \geqq 0,\ i = 1, \ldots, k$．さらに T の中で $\overrightarrow{v} \to \overrightarrow{0}$ とできるので，$c \geqq 0$．よって $\exists\overrightarrow{p} \in \mathbb{R}^k_+ \setminus \{\overrightarrow{0}\};\ \ (\forall\overrightarrow{u} \in A)\ \ (\overrightarrow{p}, \overrightarrow{u}) \geqq 0.$ 特に任意の $\overrightarrow{x} \in S$ に対して $\overrightarrow{h}(\overrightarrow{x}) \in A$ だから定理の結論を得る．（証明終わり）

鞍点定理の証明

(ii)→(i) は容易．実際 (ii) の右の不等式で $\overrightarrow{y} = \overrightarrow{0}$ とおくと，$\overrightarrow{b} \geqq \overrightarrow{0}$ と $\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$ から，$b_i\, g_i(\overrightarrow{a}) = 0,\ i = 1, \ldots, m$．よって，(iii) が成り立ち，さらに (ii) 左の不等式から $f(\overrightarrow{x}) \geqq f(\overrightarrow{x}) + (\overrightarrow{b}, \overrightarrow{g}(\overrightarrow{x})) \geqq f(\overrightarrow{a}) + (\overrightarrow{b}, \overrightarrow{g}(\overrightarrow{a})) = f(\overrightarrow{a}),\ \overrightarrow{x} \in S,\ \overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}$, だから (i) が成り立つ．

分離定理は (i)→(ii) の証明で使う．(i) が成り立つとして，$\overrightarrow{h}(\overrightarrow{x}) = \begin{pmatrix} f(\overrightarrow{x}) - f(\overrightarrow{a}) \\ \overrightarrow{g}(\overrightarrow{x}) \end{pmatrix}$ で $\overrightarrow{h} :\ S \to \mathbb{R}^{m+1}$ を定義すると，$\overrightarrow{x} \in S$ ごとに

$$\max_{i=2,\ldots,m+1} h_i(\overrightarrow{x}) = \max_{i=1,\ldots,m} g_i(\overrightarrow{x}) > 0 \text{ または } \overrightarrow{g}(\overrightarrow{x}) \leqq \overrightarrow{0}$$

であって，後者のときは (i) から $h_1(\overrightarrow{x}) = f(\overrightarrow{x}) - f(\overrightarrow{a}) \geqq 0$ だから $k = m + 1$ として値域における分離定理の仮定が成り立つので定理の結論から

$$\exists \overrightarrow{p} \in \mathbb{R}_+^{m+1} \setminus \{\overrightarrow{0}\};\ \min_{\overrightarrow{x} \in S} (\overrightarrow{p}, \overrightarrow{h}(\overrightarrow{x})) \geqq 0$$

が成り立つ．

鞍点定理の証明（続）

もし，$p_1 = 0$ ならば，p_2 から p_{m+1} までのどれかは正で他は非負だから，仮定（Slater の制約想定の十分条件）から

$(\overrightarrow{p}, \overrightarrow{h}(\overrightarrow{x}_0)) = \displaystyle\sum_{i=2}^{m+1} p_i g_{i-1}(\overrightarrow{x}_0) < 0$ となって矛盾するから，$p_1 > 0$．そこで，

$\overrightarrow{b} = \dfrac{1}{p_1}{}^t(p_2\ p_3\ \cdots\ p_{m+1}) \in \mathbb{R}_+^m$ とおけば $\overrightarrow{x} \in S$ と $\overrightarrow{y} \geqq \overrightarrow{0}$ に対して

$$\begin{aligned} L(\overrightarrow{x}, \overrightarrow{b}) &= f(\overrightarrow{a}) + f(\overrightarrow{x}) - f(\overrightarrow{a}) + (\overrightarrow{b}, \overrightarrow{g}(\overrightarrow{x})) \\ &= f(\overrightarrow{a}) + \frac{1}{p_1}(\overrightarrow{p}, \overrightarrow{h}(\overrightarrow{x})) \geqq f(\overrightarrow{a}). \end{aligned}$$

特に $\overrightarrow{x} = \overrightarrow{a}$ とおくと $\overrightarrow{b} \geqq \overrightarrow{0}$ かつ $\overrightarrow{g}(\overrightarrow{a}) \leqq \overrightarrow{0}$ だから $(\overrightarrow{b}, \overrightarrow{g}(\overrightarrow{a})) = 0$ となり，さらに (iii) が成り立つ．
よってさらに，

$$\begin{aligned} &L(\overrightarrow{x}, \overrightarrow{b}) \geqq L(\overrightarrow{a}, \overrightarrow{b}) = f(\overrightarrow{a}) \geqq f(\overrightarrow{a}) + (\overrightarrow{y}, \overrightarrow{g}(\overrightarrow{a})) = L(\overrightarrow{a}, \overrightarrow{y}), \\ &\quad \overrightarrow{x} \in S,\ \overrightarrow{y} \geqq \overrightarrow{0}. \end{aligned}$$

すなわち (ii) が成り立つ．(証明終わり)

ミニマックス定理

（非協力ゲームの均衡点の存在は不動点定理を用いる証明が多いようだが，二人零和の場合は一方の利得が他方の利得の逆符号なので，「値域での分離定理」から初等的に証明できるミニマックス定理の特別な場合.）

$\Delta_m = \{\overrightarrow{p} = (p_1, \ldots, p_m) \in \mathbb{R}_+^m \mid \sum_{j=1}^{m} p_i = 1\}$ **とおき，**

$S \subset \mathbb{R}^n$ **を凸集合，** $\overrightarrow{K}: S \to \mathbb{R}^m$ **を凸関数とする.**

$E(\overrightarrow{p}, \overrightarrow{q}) = \sum_{i=1}^{m} p_i K_i(\overrightarrow{q})$ **で** $E: \Delta_m \times S \to \mathbb{R}$ **を定義する.**

ミニマックス定理. $\min_{\overrightarrow{q} \in S} E(\overrightarrow{p}, \overrightarrow{q}) > -\infty$, $\overrightarrow{p} \in \Delta_m$, **ならば**

$$\begin{aligned} \max_{\overrightarrow{p} \in \Delta_m} \min_{\overrightarrow{q} \in S} E(\overrightarrow{p}, \overrightarrow{q}) &= \min_{\overrightarrow{q} \in S} \max_{\overrightarrow{p} \in \Delta_m} E(\overrightarrow{p}, \overrightarrow{q}) \\ &= \min_{\overrightarrow{q} \in S} E(\overrightarrow{p}^*, \overrightarrow{q}) = \max_{\overrightarrow{p} \in \Delta_m} E(\overrightarrow{p}, \overrightarrow{q}^*) = E(\overrightarrow{p}^*, \overrightarrow{q}^*) \end{aligned}$$

を満たす $(\overrightarrow{p}^*, \overrightarrow{q}^*) \in \Delta_m \times S$ **がある.**

ミニマックス定理（注）

・結論の１行目がミニマックス定理，２行目は既出の鞍点定理
・**値域の分離定理**（→鞍点定理）から得るのは
不等式 $(\min\max \leq \max\min)$
・しかし $\min\max \geq \max\min$ **は無条件**に成り立つので，
ミニマックスの形では**等号** $\min\max = \max\min$ を得る．

詳細は，服部哲弥ホームページ／講義／確率論入門
の講義スライドを参照

線形空間と「第３象限」の分離定理

・分離定理は不等号 ($\overrightarrow{p} \geqq \overrightarrow{0}$) だが等号を得る原理（線形空間との分離定理）があると等号

Gordan の定理は本講義経済数学２で既習（というよりも，主テーマ）

「親戚筋」の定理：

Gordan の定理

Stiemke の補題

Farkas の補題

・結論で（したがって仮定で）等号を許すか許さないかの違い（部分的に許すのを含めて３とおり，教科書ごとに違う定義と定理の組を採用）

・数理ファイナンスの第１基本定理

詳細は，服部哲弥ホームページ／講義／ファイナンス数学

http://web.econ.keio.ac.jp/staff/hattori/finance.htm

の講義スライドを参照

補足とまとめ

〇個別には不等号でも集合としての等号は成立

・統計的決定論におけるベイズ解集合と許容解集合の一致についてのワルドの定理（確率論入門２）

詳細は，服部哲弥ホームページ／講義／確率論入門

http://web.econ.keio.ac.jp/staff/hattori/probe.htm の秋学期講義スライドを参照

・経済学の中のいろんな分野の理論の数学は共通するものがある

・経済現象が複雑なので，通用する一般的な数学原理が限られている，とも言えよう

・分離定理はそのような,（高校で習わなかった）数学的基礎原理の１つ

文献

標準のテキスト（水色）「経済数学」慶應義塾大学経済学部

（春学期　経済数学I　第１－４章）

秋学期　経済数学II　春学期の続き　第５－７章

http://web.econ.keio.ac.jp/staff/hattori/2nen.htm

Google検索キーワード　　服部哲弥