

- 連立1次方程式は、掃き出し法で解くことができる。
- 連立代数方程式についても、掃き出し法のようなものがある。
- つるかめ算などが連立1次方程式で「頭を使わずに」解けるようなことが、連立代数方程式 + 大学数学のいくつかに見ることができる。

1 掃き出し法

(行) 階段行列：掃き出し法 (Gaussian elimination) とは、3つの行基本変形：

- ある行に、0でない数をかける
- ある行にある数 (0でもよい) をかけて、他の行に加える
- 2つの行を入れ替える

によって、与えられた行列を階段行列 (echelon matrix)：

$$\begin{pmatrix} 0 & \cdots & 0 & a_{1,j_1} & \cdots & & \cdots & * \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & a_{2,j_2} & \cdots & \vdots \\ \vdots & & & & & & & & \ddots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & a_{r,j_r} & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

に変形することである (ここで $j_1 < j_2 < \cdots < j_r$ かつ $a_{1,j_1} a_{2,j_2} \cdots a_{r,j_r} \neq 0$)。ここで非ゼロな数 $a_{1,j_1}, a_{2,j_2}, \dots, a_{r,j_r}$ は **pivot** (枢軸) と呼ばれる。

すなわち、階段行列 A とは次の様なものである：ある r が存在して (実は $r = \text{rank}(A)$ である)

1. $(r+1)$ 行目から最終行までは、すべて 0
2. 1 行目から r 行目の各行には、pivot が 1 つずつ存在して
 - pivot は「だんだん右に」分布する
 - どの pivot も「その左と下」の数 (ないかもしれない) は 0 になっている

連立方程式の解法：掃き出し法によって、連立一次方程式を解くことができる。そのあらすじは以下のとおりである：

1. 連立方程式は、 $A\mathbf{x} = \mathbf{b}$ の形に変形できる (ここで A は $m \times n$ 行列、 $\mathbf{x}$ は n 次元ベクトル、 $\mathbf{b}$ は m 次元ベクトル)。この A を係数行列、 $m \times (n+1)$ 行列 $(A, \mathbf{b})$ を拡大係数行列と呼ぶ。
2. 拡大係数行列に掃き出し法を適用し、階段行列 $(B, \mathbf{c})$ を得たとする (ここで B は $m \times n$ 行列、 $\mathbf{c}$ は m 次元ベクトル) と、与えられた連立方程式 $A\mathbf{x} = \mathbf{b}$ は $B\mathbf{x} = \mathbf{c}$ と同値になる。
3. $B\mathbf{x} = \mathbf{c}$ は、 B が階段行列なので逆向きに解くことができる (後退代入)。
4. できない場合、解は存在しない。ちなみに解が存在する必要十分条件は $\text{rank } A = \text{rank}(A, \mathbf{b})$ で、解が存在する場合の解の自由度 (パラメータの数) は $n - \text{rank } A$ である。

2 Gröbner 基底 (以下 $\mathbb{F}$ を体とし, n 変数多項式環 $\mathbb{F}[x_1, \dots, x_n]$ を考える)

零点集合: 部分集合 $S \subseteq \mathbb{F}[x_1, \dots, x_n]$ について

$$V(S) := \{(x_1, \dots, x_n) \in \mathbb{F}^n \mid \forall f \in S, f(x_1, \dots, x_n) = 0\}.$$

イデアル: 空でない部分集合 $I \subseteq \mathbb{F}[x_1, \dots, x_n]$ がイデアルであるとは

1. $\forall f_1 \in I, \forall f_2 \in I, f_1 + f_2 \in I,$
2. $\forall a \in \mathbb{F}[x_1, \dots, x_n], \forall f \in I, af \in I.$

命題: 部分集合 $S \subseteq \mathbb{F}[x_1, \dots, x_n]$ について

$$\langle S \rangle := \left\{ \sum_{f \in S} a_f f \mid a_f \in \mathbb{F}[x_1, \dots, x_n], |\{a_f \neq 0 \mid f \in S\}| < \infty \right\}$$

は S を含む最小のイデアルである. さらに $V(S) = V(\langle S \rangle).$

項順序: $\mathbb{Z}_{\geq 0}^n$ 上の全順序が項順序であるとは, 以下の 2 条件を満たすことをいう.

1. $\forall \alpha, \forall \beta, \forall \gamma \in \mathbb{Z}_{\geq 0}^n, \alpha \geq \beta \Rightarrow \alpha + \gamma \geq \beta + \gamma,$
2. $\forall \alpha \in \mathbb{Z}_{\geq 0}^n, \alpha \geq \mathbf{0}$

例: $\mathbb{Z}_{\geq 0}^n$ の辞書式順序 $\geq_{\text{lex}}$ は項順序である. ここで $\alpha, \beta \in \mathbb{Z}_{\geq 0}^n$ について, $\alpha \geq_{\text{lex}} \beta$ とは:

$$\alpha = \beta \text{ または } \lceil 1 \leq \exists j \leq n, (\alpha_j > \beta_j \text{ and } 1 \leq \forall i < j, \alpha_i = \beta_i) \rceil$$

多重次数, リーディング係数, リーディング単項式, リーディング項: $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_{\geq 0}^n$ は, $\mathbb{F}[x_1, \dots, x_n]$ の単項式 $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ と同一視できるので, $(0 \neq) f = \sum_{\alpha} f_{\alpha} x^{\alpha} \in \mathbb{F}[x_1, \dots, x_n]$ を「順序づけて」書き下すことができる. そこで非零な多項式 $f \in \mathbb{F}[x_1, \dots, x_n]$ について

1. $\text{multideg}(f) = \max\{\alpha \in \mathbb{Z}_{\geq 0}^n \mid f_{\alpha} \neq 0\},$
2. $\text{LC}(f) = f_{\text{multideg}(f)},$
3. $\text{LM}(f) = x^{\text{multideg}(f)},$
4. $\text{LT}(f) = \text{LC}(f) \text{LM}(f).$

割り算: 任意の $g_1, \dots, g_m \in \mathbb{F}[x_1, \dots, x_n] \setminus \{0\}$ と任意の $f \in \mathbb{F}[x_1, \dots, x_n]$ について, 以下をみたす $h_1, \dots, h_m, r \in \mathbb{F}[x_1, \dots, x_n]$ が存在する (この r の 1 つを $r = \text{red}(f; g_1, \dots, g_m)$ と書く. $g_1, \dots, g_m$ が Gröbner 基底のとき, r は一意的であることが証明できる).

1. $f = g_1 h_1 + \cdots + g_m h_m + r,$
2. 任意の $1 \leq i \leq m$ について, $h_i = 0$ または $\text{multideg}(g_i h_i) \leq \text{multideg}(f),$
3. $r = 0$ または $r = \sum_{\alpha} r_{\alpha} x^{\alpha}$ において $r_{\alpha} \neq 0$ ならば $1 \leq \forall i \leq m, \text{LM}(g_i) \nmid x^{\alpha}.$

Gröbner 基底の定義 : $f_1, \dots, f_m \in \mathbb{F}[x_1, \dots, x_n] \setminus \{0\}$ が Gröbner 基底とは

$$\langle \text{LT}(f) \mid f \in \langle f_1, \dots, f_m \rangle \setminus \{0\} \rangle = \langle \text{LT}(f_1), \dots, \text{LT}(f_m) \rangle.$$

S 多項式 : 非零な多項式 $f, g \in \mathbb{F}[x_1, \dots, x_n]$ の S 多項式とは

$$S(f, g) = \frac{\text{LCM}(\text{LM}(f), \text{LM}(g))}{\text{LT}(f)} f - \frac{\text{LCM}(\text{LM}(f), \text{LM}(g))}{\text{LT}(g)} g.$$

Buchberger のアルゴリズム : 与えられた非零な多項式 $g_1, \dots, g_m \in \mathbb{F}[x_1, \dots, x_n]$ について,

$\text{red}(S(g_i, g_j); g_1, \dots, g_m) \neq 0$ となる $1 \leq i < j \leq m$ が存在するとき, 1 組選んで $g_{m+1} = \text{red}(S(g_i, g_j); g_1, \dots, g_m)$ とする

という操作を繰り返す. 有限回でこの操作は終了し, $g_1, \dots, g_m$ を延長した Gröbner 基底 $g_1, \dots, g_k$ がえられる (当たり前だが, このとき $\langle g_1, \dots, g_m \rangle = \langle g_1, \dots, g_k \rangle$ となる).

消去法 : 項順序として $\geq_{\text{lex}}$ を採用する. $G = \{f_1, \dots, f_m\} \subseteq \mathbb{F}[x_1, \dots, x_n] \setminus \{0\}$ が Gröbner 基底のとき, 任意の $1 \leq k \leq n$ について $G_k := G \cap \mathbb{F}[x_k, \dots, x_n]$ も Gröbner 基底で, さらに

$$\langle G_k \rangle = \langle G_1 \rangle \cap \mathbb{F}[x_k, \dots, x_n].$$

3 Lagrange 未定乗数法

陰関数定理 : $\mathbb{R}^2$ の開集合 U で定義された C^1 級関数 $f : U \rightarrow \mathbb{R}$ を考える. $(a, b) \in U$ が, $f(a, b) = 0$ と $\partial_2 f(a, b) \neq 0$ を満たすとき, (a, b) を通る $f(x, y) = 0$ で定まる陰関数が存在する. より正確に言うと, ある $\delta > 0$ が存在して, 次が成り立つ :

1. 以下を満たす連続関数 $\varphi : (a - \delta, a + \delta) \rightarrow \mathbb{R}$ がただ一つ存在する :

- $\forall x \in (a - \delta, a + \delta), (x, \varphi(x)) \in U,$
- $\forall x \in (a - \delta, a + \delta), f(x, \varphi(x)) = 0,$
- $\varphi(a) = b.$

2. さらに φ は C^1 級で, 微分係数について以下が成り立つ

$$\forall x \in (a - \delta, a + \delta), \varphi'(x) = -\partial_1 f(x, \varphi(x)) / \partial_2 f(x, \varphi(x)).$$

極大 : $\mathbb{R}^2$ の開集合 U で定義された関数 $f : U \rightarrow \mathbb{R}$ を考える. $\mathbf{a} \in U$ で f が極大 (点) であるとは

$$\exists r > 0, \forall \mathbf{x} \in U(\mathbf{a}; r), f(\mathbf{a}) \geq f(\mathbf{x}).$$

(注) $U(\mathbf{a}; r) = \{\mathbf{x} \in \mathbb{R}^2 \mid |\mathbf{x} - \mathbf{a}| < r\}$ は, 中心 $\mathbf{a}$ で半径 r の開円盤である. 極小も同様に定義される. 極大・極小になる点を極値 (点) とよぶ.

極値の候補 : $\mathbb{R}^2$ の開集合 U で定義された関数 $f : U \rightarrow \mathbb{R}$ について, $\mathbf{a} \in U$ が f の極値点になっているとする. $\text{grad } f(\mathbf{a})$ が存在するならば, $\text{grad } f(\mathbf{a}) = \mathbf{0}$ が成り立つ.

条件付き極大 : $\mathbb{R}^2$ の開集合 U で定義された関数 $f : U \rightarrow \mathbb{R}$ を考え, さらに束縛条件を与える $g : U \rightarrow \mathbb{R}$ も考える. $\mathbf{a} \in U$ が束縛条件 $g = 0$ での極大 (点) であるとは ($g(\mathbf{a}) = 0$ かつ)

$$\exists r > 0, \forall \mathbf{x} \in U(\mathbf{a}; r), (g(\mathbf{x}) = 0 \Rightarrow f(\mathbf{a}) \geq f(\mathbf{x})).$$

Lagrange 未定乗数法 : $\mathbb{R}^2$ の開集合 U で定義された関数 C^1 級の $f, g : U \rightarrow \mathbb{R}$ について

$$\mathbf{a} \in U \text{ が束縛条件 } g = 0 \text{ での極値, かつ } \text{grad } g(\mathbf{a}) \neq \mathbf{0}$$

ならば, $\exists \lambda \in \mathbb{R}, \text{grad } f(\mathbf{a}) = \lambda \text{grad } g(\mathbf{a})$. (λ は Lagrange 未定乗数 (multiplier) とよばれる)

(注) 束縛条件下で極値を求めようとするならば, $g(x, y) = 0$ から x または y を消去して, 問題を「小さく」しようとするのが自然な考えである. しかしラグランジュ未定乗数法では, いったん変数 λ を足して, 問題を難しくしているようで興味深い (「量子力学的な」正当化もあるそうです).

4 応用

周長が一定 (簡単のため 2 とする) の三角形のうち, 面積が最大になるものは正三角形であることを示したい. 三角形の 3 辺の長さを a, b, c とすると, $a + b + c = 2$ という束縛条件のもと (本当はとりあえずさらに $0 < a, b, c < 2$), $S^2 = (1-a)(1-b)(1-c)$ (S は三角形の面積でヘロンの公式を用いた) を最大化する問題と翻訳される. ラグランジュの未定乗数法より, 連立方程式

$$a + b + c = 2, \quad -(1-b)(1-c) = \lambda, \quad -(1-a)(1-c) = \lambda, \quad -(1-a)(1-b) = \lambda$$

を解く. これを解くのは易しいが, 今は「頭を使わずに」解くことに興味がある. そこで $x_1 = a, x_2 = b, x_3 = c, x_4 = \lambda$ とし, $\geq_{\text{lex}}$ を項順序とし, $g_1 = a + b + c - 2, g_2 = bc - b - c + \lambda + 1, g_3 = ab - a - b + \lambda + 1, g_4 = ac - a - c + \lambda + 1$ に Buchberger のアルゴリズムを適用する.

1. $\text{red}(S(g_1, g_3); g_1, g_2, g_3, g_4) = b^2 - b - 2\lambda$ とできるので $g_5 := b^2 - b - 2\lambda$.
2. $\text{red}(S(g_1, g_4); g_1, g_2, g_3, g_4, g_5) = c^2 - c - 2\lambda$ とできるので $g_6 := c^2 - c - 2\lambda$.
3. $\text{red}(S(g_1, g_5); g_1, g_2, g_3, g_4, g_5, g_6) = -b\lambda - 2c\lambda + 2\lambda$ とできるので $g_7 := -b\lambda - 2c\lambda + 2\lambda$.
4. $\text{red}(S(g_1, g_6); g_1, g_2, g_3, g_4, g_5, g_6, g_7) = 3c\lambda - 2\lambda$ とできるので $g_8 := 3c\lambda - 2\lambda$.
5. $\text{red}(S(g_1, g_7); g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8) = -3\lambda^2 - \lambda/3$ とできるので $g_9 := -3\lambda^2 - \lambda/3$.

以上で繰り返しは終了し, $G = \{g_1, \dots, g_9\}$ は $\langle G \rangle = \langle g_1, \dots, g_4 \rangle$ となる Gröbner 基底である.

1. まずは λ を求める. λ の候補は $G \cap \mathbb{R}[x_4] = \{g_9\}$ より, $g_9 = 0$ を解いて $\lambda = 0, -1/9$.
2. 次に c を求める. c の候補は $G \cap \mathbb{R}[x_3, x_4] = \{g_6, g_8, g_9\}$ より, $(c, \lambda) = (0, 0), (1, 0), (2/3, -1/9)$.
3. 次に b を求める. b の候補は $G \cap \mathbb{R}[x_2, x_3, x_4] = \{g_2, g_5, g_6, g_7, g_8, g_9\}$ より, $(b, c, \lambda) = (1, 0, 0), (0, 1, 0), (1, 1, 0), (2/3, 2/3, -1/9)$.
4. 以上から $(a, b, c, \lambda) = (1, 1, 0, 0), (1, 0, 1, 0), (0, 1, 1, 0), (2/3, 2/3, 2/3, -1/9)$ と求まった.