計算代数統計

中村 文士

September 25, 2024

目次

目次

進め方

- 計算代数統計の本を読む形で進めていって、不明点があれば、この資料の中で導出していきます。
- この本は、線形代数や微積の知識すらなくとも理解できるように書かれていると思うので、他の知識の前提がなく理解できるかと思います。
- 一通り本の導出部分も追ったつもりですが、証明は説明できるほど理解できていないと思うので、適宜ご指摘頂ければと思います。

グレブナー基底でできる事の理解

今回読む本や論文、雑誌などを見て、どういった応用があるか列挙しています。

- 多項式の連立方程式 (線形でなく、x^n などの項も含む) を簡単にする時に使える
- 実験計画法の考えるべき入力の組み合わせを列挙できる
- 制御の問題で複数の多項式制約が与えられたとき、それを簡単にする時に使える
- 整数格子状の三角形分割で使える (数理科学)
- マルコフ基底というものを考える時に使う?(数理科学, Algebraic Statistics)
- transfomer を使ったある問題で使える, IBIS2023 の発表にあった

目次

目次

線形な連立方程式

以下の問題をまず考える

$$\begin{cases} x + 2y - z = 2 \\ x + y - 4z = 3 \\ x + 3y + 3z = 0 \end{cases} \tag{1}$$

これは、連立 1 次方程式なので、線形代数で学ぶ手法を使えば良い

次に、以下の問題も考えてみる

$$\begin{cases} x^2 + y^2 + 4z^2 = 81 \\ x - y + z^2 = 13 \\ xz - 2y = 18 \end{cases} \tag{2}$$

これは、第 2 式の 2 倍を第 3 式から引いて y を消去し、$xz - 2x - 2z^2 = -8$ を得て、この因数分解 $(z-2)(x-2z-4) = 0$ を行うことで、計算ができる。

次の問題は少し難しい

$$\begin{cases} x^2 + y^2 + 4z^2 = 90 \\ x - y + z = 12 \\ xz - 3y = 28 \end{cases} \tag{3}$$

これは、同じような因数分解ができない。グレブナー基底の計算のような方法で解いてみようと思ったが、教科書と同じ式導出になりそうなので、教科書で説明

注意点

- 先ほどの手順はグレブナー基底の計算アルゴリズムになっていることが、後の章で分かる。
- グレブナー基底は、各単項式 (定義はいずれ出てくる) の順序を与えると得られるもので、今回の手順は、$x \succ y \succ z$ で純辞書式順序という順序を用いた場合の結果になっている
- 最終的に得られた連立方程式の形式は z だけの式だった。これは偶然ではなく、消去定理の結果から、純辞書式順序を用いると、連立方程式のうちの一つが一つの変数の方程式になるためである (これはすごいと思った)。

実演

ここまでの式のグレブナー基底を Python 求めてみる:
Python でグレブナー基底を使う方法に例えば sympy があるので、それを簡単に紹介する

目次

いくつかの概念の定義

いくつかの概念を定義する

- 単項式: $\prod_{i=1}^{n} x_i^{a_i} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, 各 a_i は非負整数
- 単項式の次数: $\sum_{i=1}^{n} a_i$
- 多項式: 変数 $x_1, \ldots, x_n$ の有限個の項の和
- 多項式の次数: 多項式の単項式 f の次数の中で最大のもの. $deg(f)$ と表記する

※ 負の整数も含んだ多項式はローラン多項式という。この本では取り扱わないが、発展的な話題として存在しているようです。

定義したものの例

$$f = -5x_1^2x_2x_3^2 + \frac{2}{3}x_2x_4^3x_5^2 - x_3^3 - 7$$

を考える。この場合、単項式は

- $-5x_1^2x_2x_3^2$
- $\frac{2}{3}x_2x_4^3x_5^2$
- $-x_3^3$
- -7

であり、$deg(f) = 6$ である。
-7 は、係数が -7 で 1 の単項式であり、各変数の次数が 0 なので次数 0 の単項式を表す。

多項式環

K を $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ として用いる.
→ 有理数, 複素数, 実数は零割り以外の四則演算ができて、そういった集合を体と呼ぶ.
K のいずれかを係数とした $x_1, \ldots, x_n$ の多項式の集合を $K[x_1, \ldots, x_n]$ で定義する.

コメント

本を読んだとき $x_1, \ldots, x_n$ が何の集合か不明に感じましたが、不定元と呼ばれるもので、K 上の元の代入はできるが、K 上の元である必要はないもののようです。

$K[x_1, \ldots, x_n]$ の 2 つの元 f, g に対して加減算 $f + g$, $f - g$、積 fg は $K[x_1, \ldots, x_n]$ の元であるが、商 f/g は多項式とは限らないため、$K[x_1, \ldots, x_n]$ の元とはいえない。このように加減算と積について閉じている集合を環という。
$K[x_1, \ldots, x_n]$ は多項式の環なので、多項式環と呼ぶ。

アフィン多様体

イデアルを考える動機付けを行うため、まずアフィン多様体を定義する:
$K[x_1, \ldots, x_n]$ の r 個の元による連立方程式を考える;

$$\begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ f_2(x_1, \ldots, x_n) = 0 \\ \cdots \\ f_r(x_1, \ldots, x_n) = 0 \end{cases} \tag{4}$$

$K^n := \{(a_1, \ldots, a_n) | a_1, \ldots, a_n \in K\}$ と定義する。
この時、多項式 $f(x_1, \ldots, x_n)$ の零点とは K^n の元 $(a_1, \ldots, a_n)$ で

$$f(a_1, \ldots, a_n) = 0$$

を満たすものをいう (各 x_i に a_i を代入している)。アフィン多様体は、連立方程式全ての零点を与える集合 (つまり、連立方程式の解) であり、数式で書くと、以下のようになる:

$$V(f_1, \ldots, f_r) = \{(a_1, \ldots, a_n) | f_i(a_1, \ldots, a_n) = 0, i = 1, \ldots, r\}. \tag{5}$$

アフィン多様体の具体例

前節の難しかった連立方程式
$f_1 = x^2 + y^2 + 4z^2 - 90(= 0), f_2 = x - y + z - 12(= 0), f_3 = xz - 3y - 28(= 0)$ を解くことは、アフィン多様体を求めることであり、したがって f_1, f_2, f_3 のアフィン多様体は

$$V(f_1, f_2, f_3) = \left\{ \left(\frac{7 \pm \sqrt{3}}{2}, \frac{-13 \pm 3\sqrt{3}}{2}, 2 \pm \sqrt{3} \right), \left(4 \pm \sqrt{6}, \frac{-26 \pm 6\sqrt{6}}{5}, \frac{14 \pm \sqrt{6}}{5} \right) \right\}$$

である。これを解く過程を見ることで、連立方程式を解くうえで、どういった演算ができる必要があるかを確認していく。

まず、$f_2 = 0$ を x について解いて $(x = y - z + 12)$、$f_3 = xz - 3y - 28$ にそれを代入するということを行っていた:

$$\begin{cases} f_1 = x^2 + y^2 + 4z^2 - 90 = 0 \\ f_2 = x - y + z - 12 = 0 \\ f_4 = yz - 3y - z^2 + 12z - 28 = 0 \end{cases}$$

これは

$$\begin{aligned} f_4 &= (y - z + 12)z - 3y - 28 = (y - z + 12)z - xz + xz - 3y - 28 \\ &= (y - z + 12 - x)z + xz - 3y - 28 = -zf_2 + f_3 \end{aligned}$$

となっていることから、$f_4 = f_3 - zf_2$ であることが分かる。
これがもとの連立方程式と同値なことは、$f_2 = f_3 = 0 \Rightarrow f_2 = f_4 = 0,\ f_2 = f_4 = 0 \Rightarrow f_3 = zf_2 + f_4 = 0, f_2 = 0$ であることより従う。

コメント

少し後に書いてあるが、代入操作というものは、f_3 の x を $f_2 - y + z - 12$ で置き換えて余りを f_4 としていることに他ならない。

次に $f_2 = 0$ を x について解いて $(x = y - z + 12)$、$f_1 = x^2 + y^2 + 4z^2 - 90$ に代入して、$f_4 = 0$ を yz について解いたものを代入することで、

$$\begin{cases} f_2 = x - y + z - 12 = 0 \\ f_4 = yz - 3y - z^2 + 12z - 28 = 0 \\ f_5 = 2y^2 + 18y + 3z^2 - 2 = 0 \end{cases}$$

を導いた。この f_5 を導くための演算は

$$f_1 - (x + y - z + 12)f_2 + 2f_4 = f_5$$

となっていた。

さらにここから、f_4 の yz と f_5 の $2y^2$ を打ち消すため、$2yf_4 - zf_5$ を作って、$f_4 = 0$ を yz について解いた式と $f_5 = 0$ を y^2 について解いた式を代入し、最後に全体を -1 倍することで、

$$f_6 = 2y + 5z^3 - 33z^2 + 54z + 6 = 0$$

を得た。この f_6 の求めるための演算を整理すれば、

$$\begin{aligned} f_6 &= -1 \times (2yf_4 - zf_5 + 3f_5 + (2z - 6)f_4 + 6f_4) \\ &= (-2y - 2z)f_4 + (z - 3)f_5 \end{aligned}$$

となる。f_5 が不要かどうか自明でないため、ここでは一旦残しておいて、連立方程式を

$$\begin{cases} f_2 = x - y + z - 12 = 0 \\ f_4 = yz - 3y - z^2 + 12z - 28 = 0 \\ f_5 = 2y^2 + 18y + 3z^2 - 2 = 0 \\ f_6 = 2y + 5z^3 - 33z^2 + 54z + 6 = 0 \end{cases}$$

とする。

最後に

$$f_7 = -2f_4 + (z-3)f_6 = 5z^4 - 48z^3 + 155z^2 - 180z + 38$$
$$f_8 = 2f_2 + f_6 = 2x + 5z^3 - 33z^2 + 56z - 18$$

として最終的な連立方程式を得ていた (前節で解いたときは f_5 は消えていた)。

イデアル

この連立方程式を求めるうえで、行った操作を振り返ると、代入やある項を打ち消す演算というのは、元々の多項式の多項式倍と、その足し合わせで定義されることが確認できた。そこで、もとの連立方程式を定義する多項式 $f_1, \ldots, f_r$ から、多項式倍の足し合わせにより得られるすべての多項式の集合に着目する:

$$\{h_1 f_1 + \ldots + h_r f_r | h_1, \ldots, h_r \in K[x_1, \ldots, x_n]\} := \langle f_1, \ldots, f_r \rangle \tag{6}$$

これまで求めた f_4, f_5, f_6, f_7 は $\langle f_1, f_2, f_3 \rangle$ の元であることは確認出来て、例えば f_7 は

$$
\begin{aligned}
f_7 &= -2f_4 + (z-3)f_6 = -2f_4 + (z-3)((-2y-2z)f_4 + (z-3)f_5) \\
&= (-2-(z-3)(2y+2z))f_4 + (z-3)^2 f_5 \\
&= (-2-(z-3)(2y+2z))f_4 + (z-3)^2(f_1 - (x+y-z+12)f_2 + 2f_4) \\
&= (-2+(z-3)(-2y-6))f_4 + (z-3)^2 f_1 - (z-3)^2(x+y-z+12)f_2 \\
&= (z-3)^2 f_1 + (2yz^2 - 6yz + 6z^2 - 16z - (z-3)^2(x+y-12))f_2 + (-2yz+6y-6z+16) \\
&\in \langle f_1, f_2, f_3 \rangle
\end{aligned}
$$

である (途中の式展開はちょっと怪しい。ただ、f_1, f_2, f_3 の多項式倍の和であることは分かる加と思います)。

多項式の集合 $\langle f_1, \ldots, f_r \rangle$ は連立方程式 $f_1 = \cdots = f_r = 0$ を解くときの多項式倍と足し合わせの式変形により現れうる全ての多項式からなる集合である。
$\langle f_1, \ldots, f_r \rangle$ は無限個の多項式を含むことがあるが、これが元の連立方程と同じ点を解の集合 (アフィン多様体) を持つことを確認する:
$\langle f_1, \ldots, f_r \rangle$ のアフィン多様体を $V(\langle f_1, \ldots, f_r \rangle)$ と書くことにする。この時
$V(f_1, \ldots, f_r) = V(\langle f_1, \ldots, f_r \rangle)$ であることを示す。

証明.

$V(\langle f_1, \ldots, f_r \rangle) \subset V(f_1, \ldots, f_r)$ を示す。
$a \in V(\langle f_1, \ldots, f_r \rangle)$ ならば $\sum_r h_r(a) f_r(a) = 0$ が任意の $h_r \in K[x_1, \ldots, x_n]$ で成り立つ必要があるため、$f_1(a) = \cdots = f_r(a) = 0$ であることが分かる。
次に $V(f_1, \ldots, f_r) \subset V(\langle f_1, \ldots, f_r \rangle)$ を示す。
$a \in V(f_1, \ldots, f_r)$ ならば $f_1(a) = \cdots = f_r(a) = 0$ なので、$h_1(a) f_1(a) + \cdots + h_r(a) f_r(a) = 0$ が成り立つ。
これは $a \in V\langle f_1, \ldots, f_r \rangle$ であることを示しているため、$V(f_1, \ldots, f_r) = V(\langle f_1, \ldots, f_r \rangle)$ であることが示された。 □

前ページの結果から、$\langle f_1, \ldots, f_r \rangle$ のアフィン多様体を求めることで、元の連立方程式のアフィン多様体 (連立方程式の解) が求まることが分かった。
$\langle f_1, \ldots, f_r \rangle$ は以下で定義する多項式環のイデアルの一つであり、連立方程式を求めるため、イデアルの性質を使うことができるため、$\langle f_1, \ldots, f_r \rangle$ を主に考えていく:

Definition (イデアル)

多項式環の空でない部分集合 $I \subset K[x_1, \ldots, x_n]$ がイデアルとは、以下の 2 つの条件を満たすことである。

1. $f, g \in I \to f + g \in I$ (和について I の中で閉じている)
2. $f \in I, h \in K[x_1, \ldots, x_n] \to hf \in I$ (多項式倍について I の中で閉じている)

コメント

$f, g \in I$ の $-1 \in K[x_1, \ldots, x_n]$ から $-1 \times g \in I$ で、$f + (-g) = f - g \in I$ である (差についても閉じていて、そういう前提で説明していく)

また、この定義は厳密には左イデアルと言う気がするがこの教科書ではイデアルと呼ぶ。

イデアルの例

$I = \langle f_1, \ldots, f_r \rangle \subset K[x_1, \ldots, x_n]$ がイデアルであることを、定義に従って確認する。
$f, g \in \langle f_1, \ldots, f_r \rangle$ とする。$f = \sum_{i=1}^{r} h_i f_i, g = \sum_{i=1}^{r} h'_i f_i$ とすると、$f + g = \sum_{i=1}^{r} (h_i + h'_i) f_i$ となり $f + g \in I$ である。
$f = f = \sum_{i=1}^{r} h_i f_i \in \langle f_1, \ldots, f_r \rangle, h \in K[x_1, \ldots, x_n]$ とすると、$hf = \sum_{i=1}^{r} (h_i h) f_i$ となり $hf \in I$ である。
したがって、I はイデアルである。

有限生成なイデアル

イデアル $\langle f_1, \ldots, f_r \rangle$ は、有限個の多項式 $\{f_1, \ldots, f_r\}$ から定義される自然なイデアルであるが、任意のイデアルがこのように書けるかは自明でないため、それが可能か考えていく。
それを考えるため $K[x_1, \ldots, x_n]$ の空でない部分集合 $f_\lambda | \lambda \in \Lambda$ に対する有限和

$$\sum_{\lambda \in \Lambda} h_\lambda f_\lambda, h_\lambda \in K[x_1, \ldots, x_n]$$

の全体からなる集合を考える。ここで、Λ は添字集合 (有限でも、加算でも、非加算でも良い) であり、$h_\lambda | \lambda in \Lambda$ は有限個を除いて 0 である多項式の集合である。
この集合は和、多項式倍について閉じていることは自明なのでイデアルである。
このイデアルを $f_\lambda | \lambda \in \Lambda$ が生成するイデアルと呼び

$$\langle \{ f_\lambda | \lambda \in \Lambda \} \rangle$$

と書く。

逆に、任意のイデアル $I \subset K[x_1, \ldots, x_n]$ に対して、例えば、その部分集合の元そのものを使って、$I = \langle \{f_\lambda | \lambda \in \Lambda\} \rangle$ となるような $K[x_1, \ldots, x_n]$ の部分集合 $\{f_\lambda | \lambda \in \Lambda\}$ を作ることができる。この $\{f_\lambda | \lambda \in \Lambda\}$ を I の生成元と呼ぶ。
添字集合は無限の場合も考えられるが、$\langle f_1, \ldots, f_r \rangle$ は有限個の多項式から生成されていて、そういったイデアルを有限生成なイデアルと呼ぶ。

イデアルに関する諸問題

イデアルに関する内容について、ここまでの話をまとめる。
まず、多項式の連立方程式のアフィン多様体 (連立方程式の解) は、有限生成なイデアルのアフィン多様体で与えられることを学んだ。
そのため、イデアルについて考えていくことは重要である。
また、連立方程式の解を求める以外にもイデアルの応用はあり、そういった応用を考えるうえで以下の問題を考えていく。

考える問題

- イデアル記述問題: 任意に与えられたイデアル $I \subset K[x_1, \ldots, x_n]$ が、$\langle f_1, \ldots, f_r \rangle$ といった形で書けるかどうかを判定する問題
- イデアル所属問題: イデアル I と多項式 f が与えられたとき、$f \in I$ かどうかを判定する問題

イデアル記述問題, イデアル所属問題は、連立方程式の解を求める問題には不要であるが、2章の実験計画法を考える上で必要になる。

コメント

これらを解決するうえで、グレブナー基底という概念が重要である。グレブナー基底を求めることによって、消去定理が得られて、1.1 節で計算していた連立方程式の簡約化ができるため、これらの解決を目指していくことになる。

1 変数の記述問題・所属問題

まず、1 変数の多項式環 $K[x]$ について、イデアル記述問題とイデアル所属問題がどうなるかを考える。
鍵となるのは、多項式の割り算であり、例題を使って考えていく。高校の時に習った一変数の筆算を思い出すと、$f(x)$ を $g(x)$ で割ったとき、商を $q(x)$, 余りを $r(x)$ とすると、

$$f(x) = q(x)g(x) + r(x)$$

とかけて、$r(x)$ の次数は $g(x)$ より小さく、$q(x)$ と $r(x)$ は一意に定まっていた。
具体例として以下の問題を解いてみる。

問題 1.4

$f(x) = x^4 + 2x^3 - x^2 + 4x - 1$ を $g(x) = x^2 - 3x + 1$ で割った商とあまりを求めよ

これは筆算を用いるなどすれば、$f(x) = (x^2 + 5x + 13)g(x) + 38x - 14$ となる。

先ほどの計算を $g(x)$ で次数が一番大きい項である x^2 を $g(x)+3x-1$ に置き換えていると思って、$f(x)$ 上の $g(x)$ を置き換えられなくなるまで (置き換えていない項の次数が $g(x)$ の次数より小さくなるまで) 置き換えると次のようになる:

$$\begin{aligned} f(x) &= x^2(g(x)+3x-1)+(2x^3-x^2+4x-1) \\ &= x^2g(x)+5x^3-2x^2+4x-1 \\ &= x^2g(x)+5x(g(x)+3x-1)-2x^2+4x-1 \\ &= (x^2+5x)g(x)+13x^2-x-1 \\ &= (x^2+5x)g(x)+13(g(x)+3x-1)-x-1 \\ &= (x^2+5x+13)g(x)+38x-14 \end{aligned}$$