Autour du résultant et du discriminant

Shika

Notre objectif est le suivant: Déterminer de manière calculable pour deux polynômes P, Q à coefficients dans un corps $\mathbb{K}$ si ils admettent un facteur commun non trivial. De manière équivalente, on se demande comment décider algorithmiquement si P et Q admettent une racine commune dans une extension de $\mathbb{K}$. Dans le cas de polynômes à une variable, on sait déjà le faire: l'algorithme de la division euclidienne, qu'on itère pour obtenir le PGCD, répond à la question.
On va donc se concentrer sur le cas de polynômes en plusieurs variables, en privilégiant une variable, autrement dit sur le cas des polynômes de $\mathbb{K}[X_1, \cdots, X_n][Y]$. On identifiera d'ailleurs abusivement $\mathbb{K}[X, Y]$ et $\mathbb{K}[X][Y]$ sans le préciser. Et puisque la généralité ne nous coûtera rien, voire éclaircira notre travail, on considèrera simplement R un anneau commutatif, factoriel et intègre et on travaillera sur $R[X]$. Pour revenir au cas qui nous intéresse, on posera simplement $R = \mathbb{K}[X_1, \cdots X_n]$.

On fixe tout le long du document, sauf mention explicite du contraire, que R est un anneau commutatif, factoriel et intègre et que P et Q sont les deux polynômes $P = a_n X^n + \cdots + a_1 X + a_0$ et $Q = b_m X^m + \cdots + b_1 X + b_0$ de $R[X]$.

1 Résultant

1.1 Existence d'un facteur commun

Supposons que P et Q admettent un facteur commun non trivial, qu'on notera S. On peut alors écrire $P = SA$ et $Q = SB$ avec $A \in R_{n-1}[X], B \in R_{m-1}[X]$. De là, l'équation $UP + VQ = 0$, d'inconnues U et V, admet un couple solution dans un $R_{m-1}[X] \times R_{n-1}[X]$, donné par $(U, V) = (B, -A)$. Réciproquement, si P est premier avec Q et qu'on a $U, V \in R[X]$ tels que $PU = QV$, alors $P \mid V$, donc $\deg(V) \geq \deg(P)$, et symétriquement $\deg(U) \geq \deg(Q)$. On a donc la proposition suivante:

Proposition 1.1. *Deux polynômes $P, Q \in R[X]$ ont un pgcd non trivial si et seulement si leur équation de Bézout associée*

$$PU + QV = 0$$

admet une solution non triviale $U, V \in R[X]$ avec $\deg(U) < m$ et $\deg(V) < n$

On peut observer que cette équation est en fait un système, et un système linéaire. C'est ce qui va nous permettre de l'étudier. On a $R_{n+m-1}[X] = R_{m-1}[X] \oplus X^m R_{n-1}[X]$, ce qui nous permet de définir l'endomorphisme de modules

$$\varphi_{P,Q} : \begin{array}{c} (R_{m-1}[X] \oplus X^m R_{n-1}[X]) \to R_{n+m-1}[X] \\ U + X^m V \mapsto PU + QV \end{array}$$

La matrice de φ dans la base canonique sera d'un intérêt particulier, on la nomme donc:

Définition 1.1. On appelle matrice de Sylvester de P et Q la matrice d'ordre $(n+m) \times (n+m)$

$$S_{P,Q} = \begin{pmatrix} p_n & 0 & \cdots & 0 & q_m & 0 & \cdots & 0 \\ p_{n-1} & p_n & \ddots & \vdots & \vdots & q_m & \ddots & \vdots \\ \vdots & p_{n-1} & \ddots & 0 & \vdots & & \ddots & 0 \\ \vdots & \vdots & \ddots & p_n & q_1 & & & q_m \\ p_0 & & & p_{n-1} & q_0 & \ddots & \vdots & \vdots \\ 0 & \ddots & & \vdots & 0 & \ddots & q_1 & \vdots \\ \vdots & \ddots & p_0 & \vdots & \vdots & \ddots & q_0 & q_1 \\ 0 & \cdots & 0 & p_0 & 0 & \cdots & 0 & q_0 \end{pmatrix}$$

Formellement

$$(S_{P,Q})_{i,j} = \begin{cases} p_{n+j-i} \text{ si } 1 \leq i \leq m, 1 \leq j \leq n \\ q_{j-i} \text{ si } m+1 \leq i \leq n+m, 1 \leq i \leq n+m \\ 0 \text{ sinon} \end{cases}$$

On a alors immédiatement

Proposition 1.2. *Le couple $(X, Y) \in R_{m-1}[X] \times R_{n-1}[X]$ est solution de l'équation de Bézout*

$$PX + QY = 0$$

si et seulement si

$$S_{P,Q} \begin{pmatrix} X' \\ Y' \end{pmatrix} = 0$$

où X' et Y' sont les vecteurs colonnes induits respectivement par X et Y, en écrivant leurs coefficients par indices décroissants.

et donc

Théorème 1.1. *P et Q ont un facteur commun non trivial si et seulement si $\det(S_{P,Q}) = 0$.*

Cette valeur $\det(S_{P,Q})$ est souvent utile, on lui donne donc un nom, le résultant:

Définition 1.2. On appelle résultant de deux polynômes $P, Q \in R[X]$ la valeur $(P, Q) = \det(S_{P,Q})$.

On fait maintenant un petit détour par la théorie des modules pour montrer un dernier résultat sur le résultant, qui généralise les résultats précédents:

Proposition 1.3. *Soit M un R-module libre de rang n de base $\mathcal{B} = (e_1, \cdots, e_n)$ et $(x_1, \cdots, x_n)$ un n-uplet de M, qui engendre un sous-module N. Alors pour $d = \det_{\mathcal{B}}(x_1, \cdots, x_n)$ on a $dM \subset N$.*

Preuve. On écrit $x_j = x_{1,j}e_1 + \cdots x_{n,j}e_n$ pour $x_{i,j} \in \mathbb{R}$ et on pose M la matrice des $x_{i,j}$. En notant $M_{i,j}$ les cofacteurs et en développant selon la i-ème ligne, on trouve $d = \sum_{j=1}^{n} (-1)^{i+j} x_{i,j} \det(M_{i,j})$. En sommant sur i, on obtient alors $de_k = \sum_{j=1}^{n} (-1)^{i+j} x_j \det(M_{i,j})$, ce qui conclut.

Et maintenant le théorème:

Théorème 1.2. *Il existe $U \in R_{m-1}[X], V \in R_{n-1}[X]$ tels que $PU + QV = R(P, Q)$*

Preuve. On considère donc $M = R_{n+m-1}[X]$, $\mathcal{B}$ la base canonique de M et x_i la i-ème colonne de $S_{P,Q}$. Le résultat nous dit alors que $\det_{\mathcal{B}}(x_1, \cdots, x_n) R_{n+m-1}[X] = (P, Q) R_{n+m-1}[X] \subset \text{Im}(\varphi_{P,Q}) = \{PU + QV \mid \deg(U) < m, \deg(V) < n\}$. En particulier, il existe U et V satisfaisant la condition sur les degrés tels que $R(P, Q) = PU + QV$.

1.2 Expression du résultant en fonction des racines

On veut trouver une expression du résultant en fonction des racines des deux polynômes $\alpha \prod_{i \in I} (X - \alpha_i)$ et $\beta \prod_{j \in J} (X - \beta_j)$. L'idée de la preuve va être de remplacer les racines et le coefficient dominant par des indéterminées et de montrer une identité polynomiale, qui après évaluation donnera le résultat voulu.

Plus précisément, on va se placer sur l'anneau $R[A, B, (A_i)_{i \in I}, (B_j)_{j \in J}][X]$ où $A, B, (A_i)_{i \in I}, (B_j)_{j \in J}$ sont des indéterminées, et poser, juste pour cette sous-section, $P = A \prod_i (X - A_i)$ et $Q = B \prod_j (X - B_j)$.

Dans ce contexte, le résultant (P, Q) est un polynôme de $R[A, B, (A_i)_{i \in I}, (B_j)_{j \in J}]$, résultant qui sera nul si on trouve $(i, j) \in I \times J$ tels que $A_i = B_j$, puisqu'on a alors une racine commune à P et Q. Ca nous mène au lemme suivant:

Lemme 1.1. *Pour tout couple $(i,j) \in I \times J$, la relation $(A_i - B_j) \mid (P,Q)$ est vérifiée.*

Preuve. On pose p la projection canonique de $R[A, B, (A_i)_{i\in I}, (B_j)_{j\in J}]$ dans son quotient par l'idéal engendré par $A_i - B_j$. Puisque d'un côté on a $p((P,Q)) = (p(P), p(Q))$, un morphisme d'anneau préservant les polynômes, et que de l'autre on a $(p(P), p(Q)) = 0$, par le constat qui précède le lemme, on en déduit $(A_i - B_j) \mid (P,Q)$, ce qu'on voulait.

Proposition 1.4. *Le polynôme $\prod_{i,j} (A_i - B_j)$ divise (P,Q).*

Preuve. Les polynômes $(A_i - B_j)$ sont deux à deux premiers entre eux et divisent tous (P,Q), donc leur produit aussi.

On va montrer qu'en fait, à un facteur de $R[A,B]$ près, les polynômes sont égaux, en comparant les degrés en chacune des autres indéterminées:

Lemme 1.2. *Pour tout $i \in I$, $\deg_{A_i}((P,Q)) = |J|$. Similairement $\deg_{B_j}((P,Q)) = |I|$ pour tout $j \in J$.*

Preuve. En regardant les relations coefficients-racines, on voit que pour chaque $0 \leq k \leq |I|$, le k-ème coefficient p_k sera de degré 1 en A_i. On voit aussi que dans le développement en somme sur $\mathfrak{S}_n$ du déterminant de $S_{P,Q}$, on aura au plus m facteurs qui seront des coefficients de p_k dans chaque terme de la somme, majoration atteinte pour l'identité. La preuve est analogue pour le degré en les B_j.

Reste à trouver le facteur: L'inspection du déterminant à l'aide des relations coefficient-racine, encore une fois, montre que le seul monôme de $R[A,B][(A_i)_{i\in I}, (B_j)_{j\in J}]$ de degré n^m en les A_i est celui obtenu en faisant le produit des termes diagonaux, monôme de coefficient $A^m B^n$. On conclut:

Théorème 1.3. *On a l'identité $(P,Q) = A^m B^n \prod_{i,j} (A_i - B_j)$. En évaluant, on trouve:*

$$\left(\alpha \prod_{i\in I} (X - \alpha_i), \beta \prod_{j\in J} (X - \beta_j)\right) = \alpha^m \beta^n \prod_{i,j} (\alpha_i - \beta_j)$$

2 Discriminant

Une application très naturelle du résultant en caractéristique nulle est le discriminant. On suppose donc R de caractéristique nulle:

Définition 2.1. On appelle discriminant de $P \in R[X]$ la valeur $(P) = (P, P')$.

Bien sûr, le nom n'est pas choisi au hasard, c'est une généralisation du bien connu discriminant d'un trinôme du second degré, qu'on retrouve presque en calculant $(ax^2 + bx + c)$:

$$(ax^2 + bx + c, 2ax + b) =$$

$$\begin{vmatrix} a & 0 & 0 & 0 & 0 \\ b & a & 0 & 0 & 0 \\ c & b & a & 2a & 0 \\ 0 & c & b & b & 2a \\ 0 & 0 & c & 0 & b \end{vmatrix} = a \begin{vmatrix} a & 0 & 0 & 0 \\ b & a & 2a & 0 \\ c & b & b & 2a \\ 0 & c & 0 & b \end{vmatrix} = a^2 \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = a^2 \left(a \begin{vmatrix} b & 2a \\ 0 & b \end{vmatrix} - 2a \begin{vmatrix} b & 2a \\ c & b \end{vmatrix} \right) = a^3(4ac - b^2)$$

Clairement, notre discriminant s'annule si et seulement si le discriminant usuel du lycée s'annule. En fait:

Proposition 2.1. *Le discriminant de P est nul si et seulement si P a une racine double dans une clôture algébrique.*

Preuve. Supposons qu'on puisse écrire $P = (X-\alpha)^2 U$. On a alors l'égalité

$$P' = (X-\alpha)[(X-\alpha)U]' + (X-\alpha)U = (X-\alpha)([(X-\alpha)U]' + U)$$

et P et P' ont un facteur commun, d'où $(P) == (P, P') = 0$.
Réciproquement supposons que $(P) = 0$. Alors P et P' ont un facteur commun irréductible sur R, disons I. On a alors $P = IQ$ et $I \mid P' = IQ' + I'Q$ d'où $I \mid Q$. Suit que $I^2 \mid P$, et donc qu'une racine de I dans une clôture est une racine double de P.

Proposition 2.2. *Si* $P = \alpha \prod_i (X - \alpha_i)$ *alors* $(P) = \alpha^{2n-1} \prod_{i \neq j} (\alpha_i - \alpha_j)$

Preuve. On a $P' = \alpha \sum_{k=1}^{n} \prod_{i=1, i\neq k}^{n} (X - \alpha_j)$ et donc

$$(P, P') = \alpha^{n-1} \prod_{i=1}^{n} P'(\alpha_i) = \alpha^{n-1} \prod_{i=1}^{n} \alpha \prod_{j=1, j\neq i}^{n} (\alpha_i - \alpha_j) = \alpha^{2n-1} \prod_{i \neq j} (\alpha_i - \alpha_j)$$

On liste quelques applications du résultant et du discriminant

3 Applications

3.1 Théorème de Cayley-Hamilton

3.1.1 Une preuve sur $\mathbb{C}$

Soit $\mathcal{A}$ l'ensemble des matrices de $M_n(\mathbb{C})$ à valeurs propres deux à deux distinctes. Les matrices de $\mathcal{A}$ sont diagonalisables, puisque leurs polynômes caractéristiques sont nécessairement scindés à racines simples. On a donc, pour tout $M \in \mathcal{A}$, $M = PDP^{-1}$ pour P diagonale.
Si on regarde alors le polynôme caractéristique de M, on

$$\chi_M(X) = \det(PDP^{-1} - XI_n) = \det(P(D - XI_n)P^{-1}) = \det(P)\det(P)^{-1}\det(D - XI_n) = \chi_D(X)$$

On voit dans ces égalités que pour montrer l'identité $\chi_M(M) = 0$, il suffit de montrer $\chi_D(M) = P\chi_D(D)P^{-1} = 0$, ce qui est clair puisque D est une matrice diagonale.
On va déduire de cette observation le théorème suivant:

Théorème 3.1 (Cayley-Hamilton sur $\mathbb{C}$)**.** *Pour tout* $M \in M_n(\mathbb{C})$, $\chi_M(M) = 0$.

L'application $f : M_n(\mathbb{C}) \to \mathbb{C}_n[X], M \mapsto \chi_M$ est continue, donc montrer qu'elle est nulle sur une partie dense de $M_n(\mathbb{C})$ suffit à montrer qu'elle est nulle. Montrons donc que $\mathcal{A}$ est dense dans $\mathbb{C}$:

On prend $M \in M_n(\mathbb{C})$, qu'on trigonalise en une matrice triangulaire supérieure T dont on note $\lambda_1, \cdots, \lambda_n$ les coefficients diagonaux, qui sont les valeurs propres de $M = PTP^{-1}$. En les faisant légèrement varier, on peut construire une matrice triangulaire T_k égale à T sauf sur la diagonale, telle que $|T_{i,i} - (T_k)_{i,i}| \leq \frac{1}{k}$ pour tout $1 \leq i \leq n$ et $T_k \in \mathcal{A}$. La matrice T_k est alors dans $\mathcal{A}$ et la suite (T_k) converge vers T. PT_kP^{-1} est encore dans $\mathcal{A}$ (les valeurs propres sont invariantes par conjugaison), donc (PT_kP^{-1}) est une suite de $\mathcal{A}$ qui converge vers M, ce qui conclut.

3.1.2 Une généralisation sur $\mathbb{K}$

On peut fortement généraliser cette preuve, puisqu'elle marche en fait sur tout corps. La clé est de définir une topologie sur $M_k(\mathbb{K})$ pour $\mathbb{K}$ un corps quelconque, donc a priori pas muni d'une topologie canonique. On définit pour cela une topologie sur $\mathbb{K}^n$ appelée topologie de Zariski:

Lemme 3.1. *L'ensemble des parties de la forme* $V(E) = \{x \in \mathbb{K}^n \mid \forall f \in E, f(x) = 0\}$ *pour* $E \subset \mathbb{K}[X_1, \cdots, X_n]$ *vérifie les axiomes d'une base de fermés sur* $\mathbb{K}^n$

Preuve. On a $V(\{0\}) = \mathbb{K}^n, V(\{1\}) = \varnothing$, et $\bigcup_{i=1}^{n} V(E_i) = V(E)$ où $E = \{f_1 \cdots f_n \mid f_i \in E_i, 1 \leq i \leq n\}$

Définition 3.1. On appelle topologie de Zariski sur $\mathbb{K}^n$ la topologie engendrée par la base de fermés $(F_E)_{E\subset\mathbb{K}[X_1,\cdots,X_n]}$ du lemme précédent.

On peut voir $M_n(\mathbb{K})$ comme $\mathbb{K}^{n^2}$, et on a alors une topologie sur $M_n(\mathbb{K})$. Pour la preuve précédente, on utilisait deux éléments cruciaux de la topologie canonique sur $M_n(\mathbb{C})$, à savoir que les polynômes y sont continus et que $\mathcal{A}$ y est dense. On va montrer que tous les polynômes sont encore continus pour la topologie de Zariski et que les ouverts de Zariski sont denses. L'écriture $\mathcal{A} = \{M \in M_n(\mathbb{R}) \mid (\chi_M) \neq 0\}$ pour χ_A vu comme un polynôme en n^2+1 variables nous permettra de conclure que $\mathcal{A}$ est ouvert, donc dense.

Proposition 3.1. *Soit $P \in K[X_1,\cdots,X_n]$ un polynôme. L'application $P : \mathbb{K}^n \to \mathbb{K}$ est continue pour la topologie de Zariski sur $\mathbb{K}^n$ et $\mathbb{K}$.*

Preuve. Soit $F \in \mathbb{K}$ un fermé, donc un ensemble fini, alors $P^{-1}(F) = \bigcup_{x\in F} V(P-x)$ est fermé.

Pour le fait que $P^{-1}(0)$ est d'intérieur vide, on va avoir besoin du lemme suivant:

Lemme 3.2. *Pour $\mathbb{K}$ un corps infini et un polynôme $P \in K[X_1,\ldots,X_n]$ non-nul, $P^{-1}(\mathbb{K}^\times)$ est infini.*

Preuve. Le cas $n = 1$ est clair. On procède par récurrence. Supposons le résultat vrai au rang $n-1$, et soit $P = Q_nX_n^k + \cdots + Q_1X_n + Q_0$ où $Q_i \in K[X_1,\cdots,X_{n-1}]$. Le complémentaire des zéros de $Q_0\cdots Q_n$ est infini. Pour chacun des points $(x_1,\cdots,x_{n-1})$ de ce complémentaire, on a un nombre fini de x_n tels que $(x_1,\cdots,x_n)$ annule P, donc une infinité qui ne l'annulent pas. On trouve ainsi une infinité de points dans $P^{-1}(\mathbb{K}^\times)$

De ça, on déduit:

Proposition 3.2. *Un ouvert non-vide de la topologie de Zariski est dense dans $\mathbb{K}^n$*

Preuve. On observe en premier lieu que quitte à plonger $\mathbb{K}$ dans un corps infini, le résultat du lemme précédent est encore vrai. Maintenant, soit $V(E)^c$ un ouvert de $\mathbb{K}^n$ et $V(E')^c$ un autre ouvert, on veut montrer $V(E)^c \cap V(E')^c \neq \varnothing$. Il suffit, pour $p_1 \in E, p_2 \in E'$ tous deux non-nuls, de montrer $V(p_1)^c \cap V(p_2)^c$ non-vide, puisque $V(p_1)^c \subset V(E)^c$ et $V(p_2)^c \subset V(E')^c$. Mais $V(p_1p_2) = V(p_1) \cup V(p_2)$, d'où $V(p_1)^c \cap V(p_2)^c = V(p_1p_2)^c$ qui est non-vide.

En combinant tout ça, on déduit:

Théorème 3.2 (Cayley-Hamilton sur $\mathbb{K}$)**.** *Pour tout $M \in M_n(\mathbb{K})$, $\chi_M(M) = 0$.*

3.2 Théorème de Bézout faible

En gardant la notation $V(P) = \{x \in \mathbb{K}^n \mid P(x) = 0\}$ pour $P \in \mathbb{K}[X_1, \cdots, X_n]$ de la sous-section précédente, on va montrer le théorème suivant:

Théorème 3.3 (Théorème de Bézout faible)**.** *Si $P, Q \in \mathbb{K}[X, Y]$ sont deux polynômes premiers entre eux, de degrés respectifs m et n, alors $(V(P) \cap V(Q)) \leq mn$.*

On peut en fait montrer beaucoup plus fort, sur un corps algébriquement clos et dans un contexte projectif, à savoir que cette intersection est de cardinal précisément mn en comptant les multiplicités, mais ça demande nettement plus d'artillerie, trop pour ce document. On commence par énoncer lemme

Lemme 3.3. *Si $P, Q \in \mathbb{K}[X][Y]$ sont de degrés respectifs m et n, alors (P, Q) est un polynôme en X de degré majoré par mn.*

Preuve. Concisement: On note k et l les degrés de respectivement P et Q en Y, et on écrit

$$(P, Q) = \sum_{\sigma \in \mathfrak{S}_{k+l}} \varepsilon(\sigma) \alpha_{1,\sigma(1)} \cdots \alpha_{k+l,\sigma(k+l)}$$

. On vérifie alors que pour $i \leq l$, si $\sigma(i) < i$ alors $\alpha_{i,\sigma(i)} = 0$ et que sinon c'est $(l + i - \sigma(i)i)$-ème de P dans $\mathbb{K}[X][Y]$. En faisant de même pour $i > l$, on trouve $b_{i-\sigma(i)}$ si $\sigma(i) > i$, et nul sinon.
On se ramène alors à majorer la somme $\sum_{i=1}^{l} (m - k + \sigma(i) - i) + \sum_{i=l+1}^{l+k} (n - i + \sigma(i))$. La somme des $\sigma(i) - i$ étant nulle, la somme se simplifie et on trouve la majoration voulue.

On peut alors déduire de ça que $Z = (V(P) \cap V(Q))$ est fini:

Proposition 3.3. *Soit $Z = V(P) \cap V(Q)$. Si P et Q sont premiers entre eux, alors Z est fini.*

Preuve. On a $R(X) = U(X, Y)P(X, Y) + V(X, Y)Q(X, Y)$, si $(x, y) \in Z$, alors $R(x) = 0$. Comme R est un polynôme en une variable, non nul, et de degré au plus mn, il y a au plus mn valeurs de x possibles. Symétriquement, on montre qu'il y a au plus mn valeurs de y possibles, donc au plus m^2n^2 valeurs dans Z.

On déduit de ça le théorème de Bézout faible. Quitte à plonger $\mathbb{K}$ dans un corps infini, il existe deux droites vectorielles D_1, D_2 de $\mathbb{K}^2$ telle qu'aucune droite ne joignant deux points de Z ne soit parallèle ni à D_1, ni à D_2. On se place alors dans le repère $(0, D_1, D_2)$.
Dans ce repère, il y a au plus un point par abscisse, sinon la droite rejoignant deux points de même abscisse serait parallèle à D_1. Puisqu'on a montré dans la preuve précédente qu'il y avait au plus mn abscisses possibles, on en déduit la majoration du cardinal de Z par mn.