ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικό και Καποδιστριακό
Πανεπιστήμιο Αθηνών

Άλγεβρα

Ενότητα: Ο αλγόριθμος Buchberger

Ευάγγελος Ράπτης

Τμήμα Μαθηματικών

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Περιεχόμενα ενότητας

4 Ο Αλγόριθμος του Buchberger

4.1 Ο Αλγόριθμος του Buchberger

Ως τώρα έχουμε συζητήσει και αποδείξει την ύπαρξη βάσης Groebner ενός ιδεώδους I. Παρακάτω θα διατυπώσουμε έναν αλγόριθμο εύρεσης βάσης Groebner, που οφείλεται στον Buchberger, μαθητή του Groebner. Υπενθυμίζουμε ότι βάση Groebner ενός ιδεώδους I του δακτυλίου πολυωνύμων $\mathbb{F}[x_1, x_2, \cdots, x_\nu]$ είναι ένα πεπερασμένο σύνολο πολυωνύμων του I.

Μία βάση Groebner του ιδεώδους I έχει πολλές χρήσιμες ιδιότητες, μεταξύ άλλων και την ιδιότητα να παράγουν το I. Οι βάσεις Groebner ορίσθηκαν αρχικά το 1965 από τον μαθητή του W. Groebner (1899-1980), τον B.Buchberger. Επίσης ο H.Hironaka το 1965 μελετώντας δακτυλίους δυναμοσειρών ανακάλυψε ανεξάρτητα από τον B.Buchberger την ίδια έννοια της βάσης Groebner. Χρειάσθηκαν μερικά χρόνια για να αναπτυχθούν και οι υπολογιστές, ώστε η θεωρία και οι εφαρμογές των βάσεων να λάβουν τη σημερινή μορφή.

1. Έστω $f_1(x_1, x_2, \cdots, x_\nu)$ και $f_2(x_1, x_2, \cdots, x_\nu)$ δύο πολυώνυμα του δακτυλίου $\mathbb{F}[x_1, x_2, \cdots, x_\nu]$. Υποθέτουμε ότι οι μεγιστοβάθμιοι όροι των παραπάνω πολυωνύμων (μαζί με τους συντελεστές τους) είναι:

$$\alpha \cdot x_1^{\kappa_1} x_2^{\kappa_2} \cdots x_\nu^{\kappa_\nu} \text{ του } f_1(x_1, x_2, \cdots, x_\nu) \text{ και}$$
$$\beta \cdot x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_\nu^{\lambda_\nu} \text{ του } f_2(x_1, x_2, \cdots, x_\nu)$$

   Υπενθυμίζουμε ότι πρέπει $\alpha \neq 0, \beta \neq 0, \kappa_i \geq 0, \lambda_i \geq 0, i = 1, 2, \cdots, \nu$

   Ορισμός 4.1.1. Ελάχιστο κοινό πολλαπλάσιο *των μονωνύμων* $x_1^{\kappa_1} x_2^{\kappa_2} \cdots x_\nu^{\kappa_\nu}$ *και* $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_\nu^{\lambda_\nu}$ *είναι το* $x_1^{\xi_1} x_2^{\xi_2} \cdots x_\nu^{\xi_\nu}$ *όπου* $\xi_i = max(\kappa_i, \lambda_i), i = 1, 2, 3, \cdots, \nu$

2. Συνεχίζουμε με ακόμη ένα ορισμό:

   Ορισμός 4.1.2. *Έστω* $f_1(x_1, x_2, \cdots, x_\nu)$ *και* $f_2(x_1, x_2, \cdots, x_\nu)$ *δύο πολυώνυμα, όπως παραπάνω.* **S-πολυώνυμο** *των* f_1, f_2 *είναι το πολυώνυμο*

$$S(f_1, f_2) = \frac{x_1^{\xi_1} x_2^{\xi_2} \cdots x_\nu^{\xi_\nu}}{\alpha \cdot x_1^{\kappa_1} x_2^{\kappa_2} \cdots x_\nu^{\kappa_\nu}} \cdot f_1 - \frac{x_1^{\xi_1} x_2^{\xi_2} \cdots x_\nu^{\xi_\nu}}{\beta \cdot x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_\nu^{\lambda_\nu}} \cdot f_2$$

3. Ένα σημαντικό θεώρημα εδώ είναι το παρακάτω:

   Θεώρημα 4.1.3. *Έστω* $\mathbb{F}[x_1, x_2, \cdots, x_\nu]$ *ο δακτύλιος των πολυωνύμων και* $I = < f_1, f_2, \cdots, f_\mu >$ *ένα ιδεώδες αυτού. Το σύνολο* $G = \{f_1, f_2, \cdots, f_\mu\}$ *είναι βάση Groebner του ιδεώδους* I*, εάν και μόνο εάν το υπόλοιπο της διαίρεσης του* $S(f_i, f_j)$ *διά του* G *είναι μηδέν για κάθε ζεύγος* $i, j, i \neq j, 1, j = 1, 2, \cdots, \nu$.

   Απόδειξη Για την απόδειξη του βασικού αυτού θεωρήματος, το οποίο θα γίνει προσεχώς, χρειαζόμαστε μερικές προτάσεις και λήμματα:

   Το παραπάνω θεώρημα οδηγεί στον αλγόριθμο του Buchberger.

   Για τη ζωή και το έργο του Buchberger δείτε εδώ.

4. **Αλγόριθμος του Buchberger**

Βήμα 1 Τοποθετούμε σε μία σειρά τα πολυώνυμα $f_1, f_2, \cdots, f_\mu$.

Βήμα 2 Υπολογίζουμε το πολυώνυμο $S(f_1, f_2)$.

Βήμα 3 Υπολογίζουμε το υπόλοιπο της διαίρεσης του $S(f_1, f_2)$ διά του συνόλου $\{f_1, f_2, \cdots, f_\mu\}$.

Βήμα 4 Εάν το προηγούμενο υπόλοιπο είναι μηδέν, τότε συνεχίζουμε με το $S(f_1, f_3)$ διαιρώντας το με το σύνολο $\{f_1, f_2, \cdots, f_\mu\}$.
Εάν όμως δεν είναι μηδέν θεωρούμε το νέο σύνολο $\{f_1, f_2, \cdots, f_\mu, \upsilon(S(f_1, f_2))\}$[1] στη θέση του παλαιού.

Βήμα 5 Συνεχίζουμε τον αλγόριθμο ελέγχοντας όλα τα $S(f_i, f_j)$ (τα υπόλοιπά τους) και προσθέτοντας στο αρχικό σύνολο και τα πολυώνυμα $\upsilon(S(f_i, f_j))$ αν χρειάζεται.

Βήμα 6 Ο αλγόριθμος τερματίζει αν σε όλους τους ελέγχους που περιγράψαμε **όλα** τα υπόλοιπα είναι μηδέν.

Δείτε στο internet τα παρακάτω για ευρύτερη μελέτη:

1. Στη διεύθυνση εδώ για σχόλια και παραδείγματα πάνω στον αλγόριθμο.
2. Στη διεύθυνση εδώ εξαιρετική εργασία για τον αλγόριθμο Buchberger από τους Κωνσταντίνα Ρουφικτού και Άγγελο Μαντζαφλάρη.
3. Στη διεύθυνση εδώ για ένα αρκετά κατατοπιστικό άρθρο.
4. Στη διεύθυνση εδώ το άρθρο από την εγκυκλοπαίδεια Wikipedia.
5. Στη διεύθυνση εδώ μελέτη για την πολυπλοκότητα του αλγορίθμου από δύο μαθητές μου, τον Γ.Πήλιουρα και τον Δ.Διώχνο.

5 Εφαρμογές στη Γεωμετρία

5.1 Τεχνητή Νοημοσύνη

Στο κεφάλαιο αυτό θα ασχοληθούμε με τη γενική μαθηματική ιδέα της *τεχνητής νοημοσύνης.*

5.2 Το Θεώρημα βάσης του Hilbert

Θεώρημα 5.2.1. *Έστω I ένα ιδεώδες του δακτυλίου πολυωνύμων $\mathbb{F}[x_1, x_2, \cdots, x_\nu]$. Υπάρχει πεπερασμένο πλήθος πολυωνύμων του $\mathbb{F}[x_1, x_2, \cdots, x_\nu]$, το οποίο παράγει το I.*

Απόδειξη: Αν το I είναι το μηδενικό ιδεώδες, δηλαδή $I = \{\mathbf{0}\}$, τότε μπορούμε να πούμε ότι το μηδενικό πολυώνυμο του $\mathbb{F}[x_1, x_2, \cdots, x_\nu]$ παράγει το I και έτσι το θεώρημα ισχύει.

[1] Με $\upsilon(S(f_1, f_2))\}$ θα συμβολίζουμε το υπόλοιπο της διαίρεσης του $S(f_1, f_2)$ με τα υπόλοιπα πολυώνυμα.

Έστω τώρα ότι $I \neq \{\mathbf{0}\}$. Όπως έχουμε αποδείξει το I έχει (τουλάχιστον) μία βάση Groebner, έστω $B = \{g_1(x_1, x_2, \cdots, x_\nu), g_2(x_1, x_2, \cdots, x_\nu), g_3(x_1, x_2, \cdots, x_\nu), \cdots, g_\kappa(x_1, x_2, \cdots, x_\nu)\}$.

Θα αποδείξουμε ότι το B παράγει το I, δηλαδή το σύνολο των πολυωνυμικών συνδυασμών του B είναι ίσο με το I.

Έστω $f(x_1, x_2, \cdots, x_\nu)$ ένα πολυώνυμο του ιδεώδους I. Κάνουμε τη διαίρεση του $f(x_1, x_2, \cdots, x_\nu)$ διά του συνόλου B. Έχουμε:

$$\begin{aligned} f(x_1, x_2, \cdots, x_\nu) \quad = \quad & h_1(x_1, x_2, \cdots, x_\nu) \cdot g_1(x_1, x_2, \cdots, x_\nu) \ + \cdots + \\ & h_\kappa(x_1, x_2, \cdots, x_\nu) \cdot g_\kappa(x_1, x_2, \cdots, x_\nu) \ + \ \Upsilon(x_1, x_2, \cdots, x_\nu) \end{aligned}$$

όπου $\Upsilon(x_1, x_2, \cdots, x_\nu)$ το υπόλοιπο της διαίρεσης.

Την τελευταία σχέση μπορούμε να την γράψουμε:

$$\begin{aligned} \Upsilon(x_1, x_2, \cdots, x_\nu) \quad = \quad & f(x_1, x_2, \cdots, x_\nu) - h_1(x_1, x_2, \cdots, x_\nu) \cdot g_1(x_1, x_2, \cdots, x_\nu) - \cdots - \\ & h_\kappa(x_1, x_2, \cdots, x_\nu) \cdot g_\kappa(x_1, x_2, \cdots, x_\nu) \end{aligned}$$

Επειδή τα πολυώνυμα $f(x_1, x_2, \cdots, x_\nu), g_1(x_1, x_2, \cdots, x_\nu), \cdots, g_\kappa(x_1, x_2, \cdots, x_\nu)$ ανήκουν στο ιδεώδες I θα ανήκει επίσης στο ιδεώδες και το πολυώνυμο $\Upsilon(x_1, x_2, \cdots, x_\nu)$.

Όμως το σύνολο $B = \{g_1(x_1, x_2, \cdots, x_\nu), g_2(x_1, x_2, \cdots, x_\nu), g_3(x_1, x_2, \cdots, x_\nu), \cdots, g_\kappa(x_1, x_2, \cdots, x_\nu)\}$ είναι μία βάση Groebner του I, άρα ο μεγιστοβάθμιος όρος του $\Upsilon(x_1, x_2, \cdots, x_\nu)$ θα ανήκει στο ιδεώδες που παράγουν οι μεγιστοβάθμιοι όροι των πολυωνύμων του συνόλου B (δες τον ορισμό της βάσης Groebner).

Υπάρχουν τώρα δύο περιπτώσεις:

1. Το πολυώνυμο $\Upsilon(x_1, x_2, \cdots, x_\nu)$ να είναι το μηδενικό πολυώνυμο, οπότε έχουμε ότι το $f(x_1, x_2, \cdots, x_\nu)$ ανήκει στο ιδεώδες που παράγεται από το σύνολο B, κάτι στο οποίο θέλαμε να καταλήξουμε.

2. Το πολυώνυμο $\Upsilon(x_1, x_2, \cdots, x_\nu)$ να μήν είναι το μηδενικό πολυώνυμο, οπότε έχει κάποιον μη-μηδενικό μεγιστοβάθμιο όρο. Ο τελευταίος θα διαιρείται από κάποιον μεγιστοβάθμιο όρο κάποιου πολυωνύμου του B. Για να δείτε ότι ισχύει ο τελευταίος ισχυρισμός, δείτε το βίντεο εδώ ή εδώ. Όμως αν συνέβαινε κάτι τέτοιο η αρχική διαίρεση θα είχε προχωρήσει και δεν θα είχαμε αυτό το υπόλοιπο. Καταλήγουμε έτσι σε άτοπο.

Πόρισμα 5.2.2. *Ένα πολυώνυμο $f(x_1, x_2, \cdots, x_\nu) \in \mathbb{F}[x_1, x_2, \cdots, x_\nu]$, ανήκει στο ιδεώδες I εάν και μόνο εάν το υπόλοιπό του από τη διαίρεση με μία βάση Groebner του I είναι μηδέν.*

Απόδειξη : Άμεση από τα προηγούμενα.

Η πρώτη εφαρμογή σχετίζεται με το ερώτημα:

Ερώτημα 1. *Μπορεί ο υπολογιστής να αποδεικνύει μαθηματικά θεωρήματα;*

5.3 Αυτόματη απόδειξη Γεωμετρικών Θεωρημάτων

Η ιδέα της αυτόματης απόδειξης είναι η παρακάτω:

Εάν οι υποθέσεις και τα συμπεράσματα ενός θεωρήματος μπορούν να διατυπωθούν με πολυώνυμα, τότε η απόδειξη του θεωρήματος συνίσταται στην εξέταση εάν κάποια πολυώνυμα βρίσκονται σε ένα ιδεώδες ή όχι.

Δείτε πληροφορίες για τον επιστημονικό κλάδο Automated theorem proving εδώ.

Παρακάτω δίνουμε ένα παράδειγμα για να φανεί η ιδέα:

Παράδειγμα 5.3.1. *Έστω Α,Β,Δ,Γ οι κορυφές ενός παραλληλογράμμου (με τη σειρά που δίδονται). Να δειχθεί ότι οι διαγώνιες ΑΔ και ΒΓ διχοτομούνται.*

Αυτόματη γεωμετρική απόδειξη

1. Θεωρούμε ένα παραλληλόγραμμο ΑΒΔΓ με ΑΒ παράλληλο του ΔΓ και ΒΔ παράλληλο του ΑΓ.
2. Οι ιδιότητες των σχημάτων στην Ευκλείδεια Γεωμετρία παραμένουν οι ίδιες αν εφαρμόσουμε σε αυτά στροφές ή μεταφορές. Έτσι μπορούμε να θεωρήσουμε ότι η κορυφή Α είναι στην αρχή των αξόνων (0,0) και η ακμή ΑΒ στον οριζόντιο άξονα με Β= $(u_1, 0)$ για κάποιο $u_1 \neq 0$.
3. Μπορούμε να θεωρούμε το u_1 ως ανεξάρτητη μεταβλητή.
4. Η κορυφή Γ= (u_2, u_3) εισάγει δύο νέες μεταβλητές u_2 και u_3. Η κορυφή Δ δεν εισάγει νέες ανεξάρτητες μεταβλητές, διότι η δήλωση ότι το ΑΒΓΔ είναι παραλληλόγραμμο, σημαίνει, μεταξύ άλλων, ότι η θέση του Δ προσδιορίζεται πλήρως από τη θέση των Α,Β και Γ.
5. Ας συμβολίσουμε με Δ= (x_1, x_2) τις συντεταγμένες του Δ.
6. Έχουμε ότι ΑΒ παράλληλος της ΔΓ, άρα $x_2 - u_3 = 0$.
7. Έχουμε ΑΓ παράλληλος της ΒΔ, άρα $\frac{u_3}{u_2} = \frac{x_2}{x_1 - u_1}$
8. Θεωρούμε δύο πολυώνυμα

$$h_1(u_1, u_2, u_3, x_1, x_2) = x_2 - u_3 \text{ και}$$
$$h_2(u_1, u_2, u_3, x_1, x_2) = (x_1 - u_1)u_3 - x_2 u_2$$

Τα πολυώνυμα αυτά ανήκουν στον δακτύλιο $\mathbb{R}[u_1, u_2, u_3, x_1, x_2]$ των πολυωνύμων πέντε μεταβλητών με συντελεστές πραγματικούς αριθμούς. Από τα παραπάνω έχουμε ότι:

Το ΑΒΔΓ είναι παραλληλόγραμμο εάν και μόνο εάν
$$h_1(u_1, u_2, u_3, x_1, x_2) = 0 \text{ και } h_1(u_1, u_2, u_3, x_1, x_2) = 0$$

9. Ας υποθέσουμε τώρα ότι το σημείο τομής των διαγωνίων είναι το Ν= (x_3, x_4). Η δήλωση ότι το Ν είναι το σημείο τομής των διαγωνίων είναι ισοδύναμη με τη δήλωση ότι οι τριάδες σημείων (Α,Ν,Δ) και (Β,Ν,Γ) αποτελούνται από συγγραμμικά σημεία.
10. Έχουμε:
    (i) Τα Α,Ν,Δ συγγραμμικά εάν και μόνο εάν $\frac{x_4}{x_3} = \frac{u_3}{x_1}$
    (ii) Τα Β,Ν,Γ συγγραμμικά εάν και μόνο εάν $\frac{x_4}{x_3 - u_1} = \frac{u_3}{u_2 - u_1}$

11. Θεωρούμε τα πολυώνυμα

$$h_3(u_1,u_2,u_3,x_1,x_2) = x_4x_1 - x_3u_3 \text{ και}$$
$$h_4(u_1,u_2,u_3,x_1,x_2) = x_4(u_2 - u_1) - (x_3 - u_1)u_3$$

12. Μπορούμε εδώ να σκεφθούμε και να αποδείξουμε εύκολα ότι οι υποθέσεις ισχύουν εάν και μόνο εάν $h_1(u_1,u_2,u_3,x_1,x_2) = 0, h_2(u_1,u_2,u_3,x_1,x_2) = 0, h_3(u_1,u_2,u_3,x_1,x_2) = 0, h_4(u_1,u_2,u_3,x_1,x_2) = 0$, δηλαδή μετατρέψαμε τις υποθέσεις του θεωρήματος σε σχέσεις πολυωνύμων.

13. Τώρα σκεφτόμαστε σχετικά με τα ζητούμενα, δηλαδή την διερεύνηση του ερωτήματος εάν οι διαγώνιοι διχοτομούνται. Έχουμε ότι

Οι διαγώνιοι διχοτομούνται εάν και μόνο εάν $AN = N\Delta$ και $BN = N\Gamma$.

Το παραπάνω είναι ισοδύναμο με τα εξής:

(i) $AN = N\Delta$ εάν και μόνο εάν $x_3^2 + x_4^2 = (x_3 - x_1)^2 + (x_4 - x_2)^2$

(ii) $BN = N\Gamma$ εάν και μόνο εάν $(x_3 - u_1)^2 + x_4^2 = (x_3 - u_2)^2 + (x_4 - u_3)^2$

14. Θεωρούμε τα πολυώνυμα:
$g_1(u_1,u_2,u_3,x_1,x_2) = -x_3^2 - x_4^2 + (x_3 - x_1)^2 + (x_4 - x_2)^2$ και
$g_2(u_1,u_2,u_3,x_1,x_2) = (x_3 - u_2)^2 + (x_4 - u_3)^2 - (x_3 - u_1)^2 - x_4^2$

15. Προφανώς τα συμπεράσματα που θέλουμε να αποδείξουμε κωδικοποιούνται στα πολυώνυμα $g_1(u_1,u_2,u_3,x_1,x_2)$ και $g_2(u_1,u_2,u_3,x_1,x_2)$ με

$$g_1(u_1,u_2,u_3,x_1,x_2) = x_1^2 - 2x_1x_3 - 2x_4x_2 + x_2^2 \text{ και}$$
$$g_2(u_1,u_2,u_3,x_1,x_2) = 2x_3u_1 - 2x_3u_2 - 2x_4u_3 - u_1^2 + u_2^2 + u_3^2$$

16. Αναλύουμε λίγο περισσότερο τα παραπάνω: Οι υποθέσεις μας είναι ότι το σχήμα ΑΒΔΓ είναι παραληλλόγραμμο και ότι οι ευθείες ΑΔ και ΒΓ τέμνονται στο Ν. Οι υποθέσεις αυτές ισοδυναμούν με τον μηδενισμό των πολυωνύμων

$$h_1(u_1,u_2,u_3,x_1,x_2)$$
$$h_2(u_1,u_2,u_3,x_1,x_2)$$
$$h_3(u_1,u_2,u_3,x_1,x_2)$$
$$h_4(u_1,u_2,u_3,x_1,x_2)$$

Εμείς θέλουμε να αποδείξουμε ότι οι διαγώνιες διχοτομούνται. Αυτό είναι ισοδύναμο με τον μηδενισμό των πολυωνύμων $g_1(u_1,u_2,u_3,x_1,x_2)$ και $g_2(u_1,u_2,u_3,x_1,x_2)$ Αν τα πολυώνυμα g_1 και g_2 ανήκουν στο ιδεώδες $< h_1,h_2,h_3,h_4 >$, τότε τα g_1 και g_2 είναι πολυωνυμικοί συνδυασμοί των $\{h_1,h_2,h_3,h_4\}$. Αν για παράδειγμα έχουμε

$$\begin{aligned} g_1(u_1,u_2,u_3,x_1,x_2) &= \xi_1(u_1,u_2,u_3,x_1,x_2)\cdot h_1(u_1,u_2,u_3,x_1,x_2) + \cdots + \\ &\quad \xi_4(u_1,u_2,u_3,x_1,x_2)\cdot h_4(u_1,u_2,u_3,x_1,x_2) \end{aligned}$$

τότε ο μηδενισμός των $\{h_1,h_2,h_3,h_4\}$ συνεπάγεται τον μηδενισμό του g_1.

17. Ο υπολογιστής μας, λοιπόν, προκειμένου να αποδείξει το θεώρημα, θα υπολογίσει μία βάση Groebner του ιδεώδους $I =< h_1,h_2,h_3,h_4 >$, και μετά το υπόλοιπο της διαίρεσης των g_1 και g_2 με τη βάση αυτή. Γνωρίζουμε ότι τα υπόλοιπα της διαίρεσης με βάση Groebner είναι μοναδικά και τα πολυώνυμα ανήκουν στο ζητούμενο ιδεώδες εάν και μόνο εάν το υπόλοιπο είναι το μηδενικό πολυώνυμο. (**Να κάνετε εσείς τον έλεγχο**).

5.4 Άσκηση

Να θεωρήσετε ένα γεωμετρικό θεώρημα της επίπεδης γεωμετρίας και να εξετάσετε εάν ένας υπολογιστής μπορεί να το αποδείξει.