The Cayley-Hamilton Theorem

Pierre-Yves Strub
16 March 2012

MAP INTERNATIONAL
SPRING SCHOOL
ON FORMALIZATION OF
MATHEMATICS 2012
SOPHIA ANTIPOLIS, FRANCE / 12-16 MARCH

Inria informatics mathematics
ihp Institut Henri Poincaré
MICROSOFT RESEARCH INRIA JOINT CENTRE

Outline

Polynomials

Matrices

The Cayley-Hamilton Theorem

Polynomials
Definitions

Normalized (no trailing 0) sequence of coefficients:

```
Record polynomial (R : ringType) := Polynomial
  {polyseq :> seq R; _ : last 1 polyseq != 0}.
```

Are coercible to sequences:

- can directly take the k^{th} element of a polynomial (`` P`_k ``), i.e. retrieve the coefficient of X^k in `p`.
- the degree of a polynomial if its size minus 1

Polynomials

Notations

Notations:

- `{poly R}` - polynomials over `R`
- `Poly s` - the polynomial built from sequence `s`
- `'X` - monomial
- `'X^n` - monomial to the power of `n`
- `a%:P` - constant polynomial
- standard notations of ssralg (`+`, `-`, `*`, `*:`)

Can be defined by extension:

`\poly_{i < n} E` is the polynomial

`(E 0)+ (E 1) *: 'X +···+ (E n) *: 'X^n`

Polynomials

Ring operations

$$\left(\sum_{i=0}^{n} \alpha_i X^i\right)\left(\sum_{i=0}^{m} \beta_i X^i\right) = \sum_{i=0}^{n+m}\left(\sum_{j\leq i} \alpha_j \beta_{i-j}\right) X^i$$

```
Definition mul_poly (p q : {poly R}) :=
  \poly_(i < (size p + size q).-1)
    (\sum_(j < i.+1) p`_j * q`_(i - j))).
```

Polynomials

Structures

The type of polynomials has been equipped with a (commutative / integral) ring structure.

All related lemmas of `ssralg` can be used.

Polynomials

Evaluation

(Right-)evaluation of polynomials:

```
Fixpoint horner_rec s x :=
  if s is a :: s'
    then horner_rec s' x * x + a
    else 0.

Definition horner p := horner_rec p.

Notation "p .[ x ]" := (horner p x).
```

Outline

Polynomials

Matrices

The Cayley-Hamilton Theorem

Matrices

Definition

A matrix of dimension $n \times m$ over `R` is
a finite function from `'I_m * 'I_n` to `R`.

```
Inductive matrix :=
  Matrix of {ffun 'I_m * 'I_n -> R}.
```

Are coercible to functions:

- coefficient extracted by using Coq application
  `A i j` is the $(i, j)^{\text{th}}$ coefficient of `A`

Matrices

Notations

Notations:

- `'M[R]_(m, n)` - matrices of size $m \times n$ over `R`
- `'M_(m, n)`, `'M[R]_n`, `'M_n` - variants
- `a%:M` - scalar matrix (aI_n)
- `\det M`, `\tr M`, `\adj M` - determinant, trace, adjugate
- `*m` - multiplication
- standard notations of ssralg (`+`, `-`, `*`, `*:`)

Can be defined by extension:

`\matrix_{i < m, j < n} E` is the matrix

of size $m \times n$ with coefficient `E i j` at (i, j)

Matrices

Operations

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}$$

```
Definition mulmx (m n p : nat)
  (A : 'M_(m, n)) (B : 'M_(n, p))
  : 'M[R]_(m, p) :=

  \matrix_(i, j) \sum_k (A i k * B k j).
```

Matrices

Structures

The type of matrices has been equipped with a group (`zmodType`) structure.

The type of square matrices has been equipped with a ring structure.

All related lemmas of `ssralg` can be used.

Matrices

Determinant and all that

Determinant, cofactors and adjugate in 3 lines:

$$\det(A) = \sum_{\sigma \in \mathfrak{S}} \epsilon(s) \prod_i A_{i\sigma(i)}$$

```
Definition determinant n (A : 'M_n) : R :=
  \sum_(s : 'S_n) (-1) ^+ s * \prod_i A i (s i).
```

Matrices

Determinant and all that

Determinant, cofactors and adjugate in 3 lines:

$$\mathrm{cofactor}(A) : (i, j) \mapsto (-1)^{i+j} \det(\mathrm{minor}_{ij} A)$$

```
Definition cofactor n A (i j : 'I_n) : R :=
  (-1) ^+ (i + j) * determinant (row' i (col' j A)).
```

Matrices

Determinant and all that

Determinant, cofactors and adjugate in 3 lines:

$$\mathrm{adj}(A) = {}^t(\mathrm{cofactor}(A)(i,j))_{ij}$$

```
Definition adjugate n (A : 'M_n) :=
  \matrix_(i, j) cofactor A j i.
```

Outline

Polynomials

Matrices

The Cayley-Hamilton Theorem

Cayley-Hamilton

Theorem (Cayley-Hamilton)

Every square matrix over a commutative ring satisfies its own characteristic polynomial.

Characteristic polynomial

A polynomial that encodes important properties of a matrices (trace, determinant, eigenvalues):

$$
\begin{aligned}
\chi_A(X) &= \det(XI_n - A) \\
&= \begin{vmatrix} (X - A_{11}) & A_{12} & \cdots & A_{1n} \\ A_{21} & (X - A_{22}) & & \vdots \\ \vdots & & \ddots & \vdots \\ A_{n1} & \dots\dots\dots\dots\dots\dots & & (X - A_{nn}) \end{vmatrix} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \epsilon(\sigma) \prod_{1 \le i \le n} (XI_n - A)_{i\sigma(i)} \\
&= \sum_{i \le n} c_i(A) X^i \in R[X]
\end{aligned}
$$

Cayley-Hamilton

An example

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

$$\begin{aligned} \det(XI_2 - A) &= X^2 - \mathrm{tr}(A) + \det(A) \\ &= X^2 - 5X - 2 \end{aligned}$$

and

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Cayley-Hamilton

Stating the theorem

We are now ready to state the theorem

SSreflect Demo

Cayley-Hamilton

An algebraic proof

Cramer

$$\det(XI_n - A)I_n = \mathrm{adj}(XI_n - A) \cdot (XI_n - A)$$

$M_n(R[X]) \simeq M_n(R)[X]$

$$\chi_A(X) = Q(X)(X - A)$$

Evaluation at A

$$\chi_A(A) = \underbrace{(Q(X)(X - A))(A)}$$

A commutes with $(X - A)(A) = 0$

$$\underbrace{Q(A)(A - A)}_{= 0}$$

Cayley-Hamilton

An algebraic proof

The proof relies on:

- Cramer Rule:

$$\mathrm{adj}(A)\ A = \det(A) I_n$$

- $M_n(R)[X]$ and $M_n(K[X])$ are isomorphic:

$$M_n(R)[X] \xrightarrow{\simeq, \phi} M_n(K[X])$$

- Properties of right-evaluation for polynomials over non-commutative rings

Cayley-Hamilton

$M_n(R[X]) \simeq M_n(R)[X]$

Any $M \in M_n(R[X])$ can be uniquely expressed as a polynomial in $M_n(R)[X]$:

$$\begin{pmatrix} X^2+2 & 2X^2+X \\ -X & 2X+1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} X^2 + \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} X + \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Expressed using the following isomorphism:

$$\phi : M \in M_n(R[X]) \mapsto \sum_{k=0}^{\infty} ((M_{ij})_k)_{ij} X^k$$

with $((M_{ij})_k)_{ij} = 0$ whenever $k > \max_{ij} \deg(M_{ij})$

Cayley-Hamilton

$M_n(R[X]) \simeq M_n(R)[X]$

Coq Demo