The Cayley-Hamilton Theorem

Pierre-Yves Strub
16 March 2012
Outline

Polynomials

Matrices

The Cayley-Hamilton Theorem
Normalized (no trailing 0) sequence of coefficients:

```lean
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:

- \( \{\text{poly } R\} \) - polynomials over \( R \)
- \( \text{Poly } s \) - 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:

\[
\text{poly}_{\{i < n\}} E \text{ is the polynomial}
\]

\[
(E \ 0)+ (E \ 1) *: 'X + \cdots + (E \ n) *: 'X^n
\]
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\}) :=
\[\text{poly (i < (size p + size q).-1)}\]
\[\text{(\sum (j < i.+1) p'_{-j} * q'_{-(i - j)})}].\]
The type of polynomials has been equipped with a (commutative / integral) ring structure.

All related lemmas of ssralg can be used.
(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
A matrix of dimension $n \times m$ over $\mathbb{R}$ is a finite function from $\mathbb{I}_m \times \mathbb{I}_n$ to $\mathbb{R}$.

Inductive matrix :=
Matrix of \{ffun \mathbb{I}_m \times \mathbb{I}_n \to \mathbb{R}\}.

Are coercible to functions:

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

Notations:

- $M_{(m, n)}^R$ - matrices of size $m \times n$ over $\mathbb{R}$
- $M_{(m, n)}$, $M_{(R, n)}$, $M_n$ - variants
- $a\cdot M$ - scalar matrix $(aI_n)$
- $\det M$, $\text{tr} M$, $\text{adj} M$ - determinant, trace, adjugate
- $\cdot m$ - multiplication
- standard notations of ssralg ($+$, $-$, $\cdot$, $\cdot :$)

Can be defined by extension:

\[
\text{matrix}_{\{i < m, j < n\}} E \text{ is the matrix of size } m \times n \text{ with coefficient } E \text{ at } (i, j)
\]
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) :=

\[\text{\textbackslash matrix}(i, j) \text{\textbackslash sum}(k \ (A i k * B k j)).\]
Matrices

Structures

The type of matrices has been equipped with a group \((\mathbb{Z}/m\mathbb{Z})\) structure.

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

All related lemmas of \texttt{ssralg} can be used.
Matrices

Determinant and all that

Determinant, cofactors and adjugate in 3 lines:

\[
\text{det}(A) = \sum_{\sigma \in S} \epsilon(s) \prod_{i} A_{i\sigma(i)}
\]

\textbf{Definition} determinant n (A : \textbf{M}_n) : \mathbb{R} :=
\sum_{s : \textbf{S}_n} (-1)^s * A_{i\sigma(i)}.
Matrices
Determinant and all that

Determinant, cofactors and adjugate in 3 lines:

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

**Definition** cofactor n A (i j : 'I_n) : R :=
(-1) ^(i + j) \cdot \operatorname{determinant} (\text{row’ } i \text{ (col’ } j \text{ A))}.
Matrices

Determinant and all that

Determinant, cofactors and adjugate in 3 lines:

\[ \text{adj}(A) = t(\text{cofactor}(A)(i,j))_{ij} \]

\textbf{Definition} \ \text{adjugate n (A : 'M_n)} := \\
\text{\textbackslash matrix}(i, j) \ \text{cofactor A j i.}
Outline

Polynomials

Matrices

The Cayley-Hamilton Theorem
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):

\[
\chi_A(X) = \det(XI_n - A) = \begin{vmatrix}
(X - A_{11}) & A_{12} & \cdots & A_{1n} \\
A_{21} & (X - A_{22}) & \cdots & \\
\vdots & \ddots & \ddots & \\
A_{n1} & \cdots & (X - A_{nn})
\end{vmatrix}
\]

\[
= \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{1 \leq i \leq n} (XI_n - A)_{i\sigma(i)}
\]

\[
= \sum_{i \leq n} c_i(A)X^i \in R[X]
\]
Cayley-Hamilton

An example

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

\[ \det(XI_2 - A) = X^2 - \text{tr}(A) + \det(A) \]

\[ = X^2 - 5X - 2 \]

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

\[ M_n(R[X]) \cong M_n(R)[X] \]

Cramer

\[ \det(XI_n - A)I_n = \text{adj}(XI_n - A) \cdot (XI_n - A) \]

Evaluation at \( A \)

\[ \chi_A(X) = Q(X)(X - A) \]

\[ \chi_A(A) = (Q(X)(X - A))(A) \]

\[ A \text{ commutes with } (X - A)(A) = 0 \]

\[ Q(A)(A - A) = 0 \]
Cayley-Hamilton
An algebraic proof

The proof relies on:

- Cramer Rule:

\[ \text{adj}(A) \cdot A = \det(A)I_n \]

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

\[ M_n(R)[X] \overset{\sim, \phi}{\longrightarrow} M_n(K[X]) \]

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

\[ M_n(R[X]) \cong M_n(R)[X] \]

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