planetmath.org

Math for the people, by the people.

lecture notes on the Cayley-Hamilton theorem

Canonical name	LectureNotesOnTheCayleyHamiltonTheorem
Date of creation	2013-03-22 16:50:59
Last modified on	2013-03-22 16:50:59
Owner	rmilson (146)
Last modified by	rmilson (146)
Numerical id	8
Author	rmilson (146)
Entry type	Topic
Classification	msc 15A18
Classification	msc 15A15

1   Overview

You should all know about the characteristic polynomial of a square matrix, A. To calculate the characteristic polynomial of A, one subtracts a variable, say t, from the diagonal entries of A, and then takes the determinant of the result. In other words, letting $p_A(t)$ denote the characteristic polynomial of A, one has

$$p_A(t) = \det(A - tI),$$

where I as usual denotes the identity matrix. For example, set

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Evaluating the determinant of $A - tI$ one gets

$$p_A(t) = -t^3 - t^2 + 6t + 7.$$

Now the interesting thing about square matrices is that one can do algebra with them. So if A is a 3×3 matrix then A^2, A^3, indeed every power of A will also be a 3×3 matrix. Indeed, one can take any polynomial $p(t)$, and happily plug A into it. The result will be some other 3×3 matrix. The obvious question now is: what will happen when one plugs a square matrix A into its own characteristic polynomial? Let's see what happens for the sample 3×3 matrix above. Straightforward calculations show that

$$A^2 = \begin{bmatrix} 4 & 1 & 5 \\ 1 & 5 & -2 \\ 1 & 0 & 4 \end{bmatrix}, \qquad A^3 = \begin{bmatrix} 9 & 11 & 13 \\ -1 & -10 & 8 \\ 5 & 6 & 3 \end{bmatrix},$$

Next adding the various powers of A with the coefficients of characteristic polynomial (note that one uses the identity matrix in place of the constants) one gets

$$\overset{\displaystyle -A^3}{} \qquad\qquad \overset{\displaystyle -A^2}{} \qquad\qquad \overset{\displaystyle 6A}{} \qquad\qquad \overset{\displaystyle 7I}{}$$

$$-\begin{bmatrix} 9 & 11 & 13 \\ -1 & -10 & 8 \\ 5 & 6 & 3 \end{bmatrix} -\begin{bmatrix} 4 & 1 & 5 \\ 1 & 5 & -2 \\ 1 & 0 & 4 \end{bmatrix} +6\begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{bmatrix} +7\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

1

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Zero! One gets zero. This seemingly miraculous answer is not a coincidence. Indeed one gets zero regardless of what matrix one starts with. I encourage you to try this with a few examples of your own.

Theorem 1 (Cayley-Hamilton). *Let A be an $n \times n$ matrix, and let $p_A(t) = \det(A - tI)$ be the corresponding characteristic polynomial. Then, $p_A(A) = 0$.*

The goal of these notes will be to explain and prove the above theorem. There are various hidden reasons that make the Cayley-Hamilton theorem work. It is the purpose of these notes to bring these reasons into the open.

2 The Gist of the Matter.

Indeed there are two factors that make the Cayley-Hamilton theorem such a striking and interesting result. Recall that if U is an n-dimensional vector space, and $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_{n+1}$, are any $n + 1$ vectors in U, then there will some kind of a linear relation between the $\mathbf{u}_i$'s, i.e. for some choice of scalars $a_1, a_2, \ldots, a_{n+1}$ one will have

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + a_3\mathbf{u}_3 + \ldots a_{n+1}\mathbf{u}_{n+1} = 0.$$

Now the space of 3×3 matrices is 9-dimensional. Therefore for every 3×3 matrix A there must be a linear relationship between the 10 different matrix powers

$$A^9, A^8, A^7, A^6, A^5, A^4, A^3, A^2, A^1 = A, A^0 = I.$$

The "miracle" of the Cayley-Hamilton theorem is twofold. First, a linear relation arises already for the powers A^3, A^2, A^1, A^0. Second, the coefficients for this linear relation are precisely the coefficients of the characteristic polynomial of A.

Let's put it another way. Look at the first column vectors of the matrices A^3, A^2, A^1, A^0, i.e. the vectors

$$\begin{bmatrix} 9 \\ -1 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Now $\mathbb{R}^3$ is a 3-dimensional vector space, and so there should be a linear relation between the above 4 vectors. Indeed there is: the coefficients of the linear relation are $-1, -1, 6, 7$ (i.e. -1 times the first vector, plus -1 times the second, plus 6 times the third, plus 7 times the fourth is equal to zero — try it yourself!). What about the second column vectors of A^3, A^2, A^1, A^0? Now the vectors in question are

$$\begin{bmatrix} 11 \\ -10 \\ 6 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Again, we have here 4 vectors from a 3-dimensional vectors space, and therefore there should be a linear relation between the vectors. However by some miracle the coefficients of the linear relation for the second column vectors are the same as the coefficients of the linear relation between the first column vectors, namely $-1, -1, 6, 7$. Furthermore, these coefficients are precisely the coefficients of the characteristic polynomial: $-t^3 - t^2 + 6t + 7$. Needless to say the third column vectors are joined in a linear relation with the same coefficients: $-1, -1, 6, 7$. Why is this happening?

3 The Cyclic Basis

Let's look again at the first column vectors of the matrices A^0, A^1, A^2 (recall that A^0 is just the identity matrix):

$$\mathbf{u}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix},$$

and let's take these 3 vectors as a new basis, $\mathbf{B}$. A basis obtained in this fashion, i.e. by starting with a vector and successively applying a matrix to it, is called a cyclic basis. What will be the representation of the matrix A relative to this cyclic basis? Now $\mathbf{u}_0$ is just the first elementary vector, $\mathbf{e}_1$. Furthermore, note that $\mathbf{u}_1$ is nothing but $A\mathbf{u}_0$, and that $\mathbf{u}_2 = A\mathbf{u}_1$. Now $A\mathbf{u}_2$ is the first column vector of A^3 and we already determined the linear relation between the first column vectors of $A^3, \ldots A^0$. The bottom line is that

$$A\mathbf{u}_2 = -(\mathbf{u}_2 - 6\mathbf{u}_1 - 7\mathbf{u}_0),$$

3

and consequently A will have the following appearance relative to the basis **B**:

$$[A]_{\mathbf{B}} = \begin{bmatrix} 0 & 0 & 7 \\ 1 & 0 & 6 \\ 0 & 1 & -1 \end{bmatrix}$$

The transition matrix P from **B** to the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is given by

$$P = \begin{bmatrix} 1 & 1 & 4 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Of course P is relevant to our discussion precisely because

$$[A]_{\mathbf{B}} = P^{-1}AP.$$

Proposition 1. *Let A be an $n \times n$ matrix, P a non-singular $n \times n$ matrix, and set $B = P^{-1}AP$. The matrices A and B have the same characteristic polynomials.*

Proof. The characteristic polynomial of B is given as

$$\det(B - tI) = \det\left(P^{-1}AP - tI\right) = \det\left(P^{-1}(A - tI)P\right).$$

Recall that the determinant of a product is the product of the determinants, and that the determinant of an inverse is the inverse of the determinant. Therefore

$$\det(B - tI) = \det(P^{-1})\det(A - tI)\det(P) = \det(A - tI).$$

$\square$

In other words, according to the above theorem we should expect the characteristic polynomial of $[A]_{\mathbf{B}}$ to be equal to the characteristic polynomial of A. Let's check this using a co-factor expansion.

$$\begin{vmatrix} -t & 0 & 7 \\ 1 & -t & 6 \\ 0 & 1 & -1-t \end{vmatrix} = -t(t^2 + t - 6) + 7 \cdot 1 = -t^3 - t^2 + 6t + 7$$

Also note that the last column of $[A]_{\mathbf{B}}$ contains all but one of the coefficients of the characteristic polynomial. This too is not a coincidence.

4

Proposition 2. *Consider an $n \times n$ matrix B such that the jth column vector for $j = 1, 2, \ldots n - 1$ is the basic vector $\mathbf{e}_{j+1}$, while the last column of B is the vector $[-b_0, -b_1, \ldots, -b_{n-1}]^T$. In other words B has the following form:*

$$
B = \begin{bmatrix}
0 & 0 & 0 & \ldots & 0 & -b_0 \\
1 & 0 & 0 & \ldots & 0 & -b_1 \\
0 & 1 & 0 & \ldots & 0 & -b_2 \\
0 & 0 & 1 & \ldots & 0 & -b_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 1 & -b_{n-1}
\end{bmatrix}
$$

Then, the characteristic polynomial of B is given by

$$(-1)^n p_B(t) = t^n + b_{n-1} t^{n-1} + \ldots + b_2 t^2 + b_1 t + b_0.$$

Proof. We will calculate the determinant of $B - tI$ by doing a co-factor expansion along the first row. Let B_1 be the matrix obtained by deleting the first row and the first column from $B - tI$, and let D be the matrix obtained by deleting the first row and the last column from $B - tI$. Doing a co-factor expansion along the top row, it is easy to see that

$$\det(B - tI) = -t \det(B_1) + (-1)^n b_0 \det(D).$$

Now D is an upper triangular matrix with ones on the diagonal, and therefore $\det(D) = 1$. The matrix B_1, on the other hand has the same structure as $B - tI$, only it's 1 size smaller. To that end let B_2 be the matrix obtained by deleting the first two rows and columns from $B - tI$. By the same reasoning as above it's easy to see that

$$\det(B_1) = -t \det(B_2) + (-1)^{n-1} b_1,$$

and therefore

$$\det(B - tI) = (-1)^n b_0 - t \Big((-1)^{n-1} b_1 - t \det(B_2) \Big).$$

Continuing inductively we see that for even n, the determinant of $B - tI$ will have the form:

$$b_0 - t \left(-b_1 - t \left(b_2 - t \left(-b_3 - t(\ldots) \right) \right) \right) = b_0 + b_1 t + b_2 t^2 + b_3 t^3 + \ldots + b_{n-1} t^{n-1} + t^n$$

For odd n, $\det(B - tI)$ will be just like the formula above, but multiplied through by a negative sign. $\qquad \square$

4   Putting it all together

Thanks to Propositions **??** and **??** we are now in a position to understand and to prove the Cayley-Hamilton Theorem. Let A be an $n \times n$ matrix. Start by setting $\mathbf{u}_0 = \mathbf{e}_1$, and then create a sequence of vectors by successively applying A, i.e. $\mathbf{u}_1 = A\mathbf{u}_0$, $\mathbf{u}_2 = A\mathbf{u}_1$, etc. Notice that $\mathbf{u}_k = A^k\mathbf{u}_0$; in other words, $\mathbf{u}_k$ is the first column of the matrix A^k.

Next, suppose that the n vectors $\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{n-1}$ form a basis, $\mathbf{B}$, of $\mathbb{R}^n$ (There are matrices A for which this doesn't happen, but we'll consider this possibility later.) There will therefore exist scalars $b_0, b_1, \ldots b_{n-1}$ such that

$$\mathbf{u}_n + b_{n-1}\mathbf{u}_{n-1} + \ldots + b_1\mathbf{u}_1 + b_0\mathbf{u}_0 = 0.$$

Now the representation of A relative to the cyclic basis $\mathbf{B}$ will have the form

$$[A]_\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & -b_0 \\ 1 & 0 & 0 & \ldots & 0 & -b_1 \\ 0 & 1 & 0 & \ldots & 0 & -b_2 \\ 0 & 0 & 1 & \ldots & 0 & -b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -b_{n-1} \end{bmatrix}$$

By Proposition **??** the characteristic polynomial of $[A]_\mathbf{B}$ is equal to the characteristic polynomial of A. Furthermore, by Proposition **??** the characteristic polynomial of $[A]_\mathbf{B}$ is equal to

$$\pm(t^n + b_{n-1}t^{n-1} + \ldots b_1 t + b_0).$$

Only one conclusion is possible: $b_0, b_1, \ldots, b_{n-1}$ must be precisely the coefficients of the characteristic polynomial of A. Let us summarize these findings.

Proposition 3. *Let A be an $n \times n$ matrix, with characteristic polynomial*

$$p_A(t) = \pm \left(t^n + b_{n-1}t^{n-1} + \ldots + b_1 t + b_0 \right).$$

Fix a number k between 1 and n, and let $\mathbf{u}_j$ be the kth column of the matrix A^j. If the vectors $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ form a basis of $\mathbb{R}^n$, then the vectors $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}, \mathbf{u}_n$ satisfy the linear relation:

$$\mathbf{u}_n + b_{n-1}\mathbf{u}_{n-1} + \ldots + b_1\mathbf{u}_1 + b_0\mathbf{u}_0 = 0.$$

5 A Complication

We are almost done with the proof of the Cayley-Hamilton Theorem. First, however, we must deal with the possibility that the square matrix A is such that the column vectors of $A^0, A^1, \ldots, A^{n-1}$ do not form a basis. Consider, for example

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 4 & -4 \\ 1 & 2 & -2 \end{bmatrix}$$

An easy calculation shows that the characteristic polynomial is given by

$$p_A(t) = t^3 - 3t^2 + t + 2.$$

Writing down the sequence of powers of A:

$$\begin{array}{cccc} A^3 & A^2 & A^1 & A^0 \end{array}$$

$$\begin{bmatrix} 3 & -2 & 4 \\ 2 & 15 & -14 \\ 2 & 7 & -6 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 & 2 \\ 1 & 7 & -6 \\ 1 & 3 & -2 \end{bmatrix} \quad \begin{bmatrix} 1 & -1 & 2 \\ 1 & 4 & -4 \\ 1 & 2 & -2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

we notice that the first columns do, in fact, obey a linear relation with the coefficients of the characteristic polynomial:

$$\begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} - 3 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{1}$$

However these first column vectors *do not* form a basis of $\mathbb{R}^3$, and therefore Proposition **??** is not enough to explain why these vectors obey the above linear relation.

In order to find an explanation, let us proceed as follows. Just as before, start by setting $\mathbf{u}_0 = \mathbf{e}_1$, and $\mathbf{u}_1 = A\mathbf{u}_0$. If we take $\mathbf{u}_2 = A\mathbf{u}_1$, then $\mathbf{B} = (\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2)$ will not form a basis, so instead, let us choose $\mathbf{u}_2$ that is linearly independent from $\mathbf{u}_0$ and $\mathbf{u}_1$, thereby ensuring that $\mathbf{B}$ is a basis. There are many, many possible such choices for $\mathbf{u}_2$. To keep the discussion concrete, let us take $\mathbf{u}_2 = \mathbf{e}_3 = [0, 0, 1]^T$. Note that

$$A\mathbf{u}_0 = \mathbf{u}_1$$

$$A\mathbf{u}_1 = [2, 1, 1]^T = \mathbf{u}_0 + \mathbf{u}_1.$$

$$A\mathbf{u}_2 = [2, -4, 2]^T = 6\mathbf{u}_0 - 4\mathbf{u}_1 + 2\mathbf{u}_2.$$

Therefore, representing A relative to the basis $\mathbf{B}$ we obtain

$$[A]_{\mathbf{B}} = \begin{bmatrix} 0 & 1 & 6 \\ 1 & 1 & -4 \\ 0 & 0 & 2 \end{bmatrix}$$

By Proposition **??**, we know that the characteristic polynomial of $[A]_{\mathbf{B}}$ is equal to the characteristic polynomial of A. However, we know much more.

Proposition 4. *Let B be an $n \times n$ matrix of the form*

$$B = \begin{bmatrix} B_1 & B_2 \\ \mathbf{0} & B_3 \end{bmatrix}$$

where B_1 is a $k \times k$ matrix, B_2 is a $k \times (n - k)$ matrix, and B_3 is a $(n - k) \times (n - k)$ matrix. Then, the characteristic polynomial of B is the product of the characteristic polynomials of B_1 and B_3, i.e. $p_B(t) = p_{B_1}(t) \times p_{B_3}(t)$.

Proof. Note that
$$B - tI = \begin{bmatrix} B_1 - tI_1 & B_2 \\ \mathbf{0} & B_3 - tI_3 \end{bmatrix}$$

where I_1 is the $k \times k$ identity matrix, and I_3 is the $(n - k) \times (n - k)$ identity matrix. The Proposition now follows from the fact that the determinant of a matrix whose shape is like B is the determinant of the upper-left block times the determinant of the lower-right block. $\square$

Thanks to Proposition **??** we know that the characteristic polynomial of $[A]_{\mathbf{B}}$ is a product of the characteristic polynomial of the 2×2 matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

and the characteristic polynomial of the 1×1 matrix $[2]$. In other words,

$$p_{[A]_{\mathbf{B}}}(t) = (t^2 - t - 1)(t - 2).$$

8

Furthermore by Proposition **??** we know that $A\mathbf{u}_1$, $\mathbf{u}_1$, $\mathbf{u}_0$, i.e. the first column vectors of A^2, A^1, A^0, obey a linear relation with the coefficients of the polynomial $t^2 - t - 1$:

$$A\mathbf{u}_1 - \mathbf{u}_1 - \mathbf{u}_0 = \mathbf{0},$$

$$\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2}$$

Multiplying this relation through by A wed deduce that the first column vectors of A^3, A^2, A^1 obey the same linear relation:

$$\begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3}$$

Next think about what it means to multiply a polynomial such as $t^2 - t - 1$ by another polynomial such as $t - 2$. Indeed, one can structure the multiplication by multiplying the first polynomial through by t, then multiplying it through by -2, and then adding the two terms:

$$\begin{array}{ccccc} t^3 & -t^2 & -t & & \\ & -2t^2 & 2t & 2 & \\ \hline t^3 & -3t^2 & +t & +2 & \end{array}$$

The bottom line is, of course, just the characteristic polynomial of A, and the whole idea behind the above calculation is that $p_A(t)$ can be "formed out of" the polynomial $t^2 - t - 1$. This shows that we can combine relations (**??**) and (**??**) and produce in the end the desired relation (**??**). All we have to do is take relation (**??**), and add to it -2 times the relation (**??**). This explains why the first column vectors of A^3, A^2, A^1, A^0 obey a linear relation whose coefficients come from the characteristic polynomial of A.

Proposition 5. *Let A be an $n \times n$ matrix, with characteristic polynomial*

$$p_A(t) = \pm \left(t^n + b_{n-1}t^{n-1} + \ldots + b_1 t + b_0 \right).$$

Fix a number k between 1 and n, and let $\mathbf{u}_j$ be the kth column of the matrix A^j. The vectors $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}, \mathbf{u}_n$ satisfy the linear relation:

$$\mathbf{u}_n + b_{n-1}\mathbf{u}_{n-1} + \ldots + b_1\mathbf{u}_1 + b_0\mathbf{u}_0 = 0,$$

even if the vectors $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ do not form a basis of $\mathbb{R}^n$.

Proof. Suppose that there is a number $m < n$ such that $\mathbf{u}_m$ can be given as a linear combination of $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{m-1}$; let's say

$$\mathbf{u}_m + c_{m-1}\mathbf{u}_{m-1} + \ldots + c_1\mathbf{u}_1 + c_0\mathbf{u}_0 = 0.$$

Choose vectors $\mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{n-1}$ so that the list

$$\mathbf{B} = (\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{m-1}, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{n-1})$$

forms a basis of $\mathbb{R}^n$. Relative to this basis, A will have the form

$$A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix},$$

where the upper-left block has the form

$$A_1 = \begin{bmatrix} 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & \ldots & 0 & -c_1 \\ 0 & 1 & \ldots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -c_{m-1} \end{bmatrix}$$

By Proposition **??**,

$$p_{A_1}(t) = \pm(t^m + c_{m-1}t^{m-1} + \ldots + c_1 t + c_0)$$

. Let $\pm(t^{n-m} + d_{n-m-1}t^{m-1} + \ldots d_1 t + d_0)$ denote the characteristic polynomial of A_3. By Proposition **??**, the characteristic polynomial of $[A]_{\mathbf{B}}$ (which is equal to the characteristic polynomial of A) is the product $p_{A_1}(t) \times p_{A_3}(t)$, and therefore $p_A(t)$ can be obtained by taking linear combinations of $p_{A_1}(t)$ times various powers of t:

$d_0 \times$			$t^m +$	$c_{m-1}t^{m-1} +$	$\ldots +$	$c_2 t^2 +$	$c_1 t +$	c_0	
$d_1 \times$		$t^{m+1} +$	$c_{m-1}t^m +$	$c_{m-2}t^{m-1} +$	$\ldots +$	$c_1 t^2 +$	$c_0 t$		
		$\ldots$		$\ldots$	$\ldots$				
$1 \times$	$t^n +$	$c_{m-1}t^{n-1} +$		$\ldots +$	$c_1 t^{n-m+1} +$	$c_0 t^{n-m}$			
	$t^n +$	$b_{n-1}t^{n-1} +$	$b_{n-2}t^{n-2} +$		$\ldots$		$+$ $b_2 t^2 +$	$b_1 t +$	b_0	

Corresponding to the above polynomials are the relations

$\mathbf{u}_m +$	$c_{m-1}\mathbf{u}_{m-1} +$	$\ldots +$	$c_2\mathbf{u}_2 +$	$c_1\mathbf{u}_1 +$	$c_0\mathbf{u}_0$	$= 0$	
$\mathbf{u}_{m+1} +$	$c_{m-1}\mathbf{u}_m +$	$c_{m-2}\mathbf{u}_{m-1} +$	$\ldots +$	$c_1\mathbf{u}_2 +$	$c_0\mathbf{u}_1$	$= 0$	
	$\ldots$	$\ldots$	$\ldots$				
$\mathbf{u}_n +$	$c_{m-1}\mathbf{u}_{n-1} +$	$\ldots +$	$c_1\mathbf{u}_{n-m+1} +$	$c_0\mathbf{u}_{n-m}$		$= 0$	

Adding these relations in the same way as the polynomials yields the desired relation:

$$\mathbf{u}_n + b_{n-1}\mathbf{u}_{n-1} + \ldots + b_1\mathbf{u}_1 + b_0\mathbf{u}_0 = 0,$$

$\square$

Now we really are finished. Thanks to Propositions **??** and **??** we know that for every $k = 1, \ldots, n$, the kth column vectors of the matrices

$$A^n, A^{n-1}, \ldots, A^1, A^0$$

obey a linear relation with the coefficients of the characteristic polynomial of A. Since this is true for every column of the above matrices, it is certainly true for the full matrix — and that is the precisely the conclusion of the Cayley-Hamilton theorem.