Matrices over $\mathbb{Z}$, $\mathbb{F}[t]$

We work with matrices over a ring R.
We assume R is commutative (which means
we can use determinants.

Matrices A & B over R are equivalent if there
are invertible matrices P and Q such that

$$B = PAQ.$$

This is an equivalence relation.

Theorem Let $\mathbb{F}$ be a field. Two $n \times n$ matrices over $\mathbb{F}$ are similar if & only if $tI-A$ and $tI-B$ are equivalent over $\mathbb{F}[t]$.

We don't prove this now.
One reason it holds is that $\mathbb{F}[t]$ is a principal ideal domain.

Transfer matrices

Consider a dynamical system

$$(1) \quad \begin{aligned} x_{n+1} &= A x_n + B u_n \\ y_n &= C x_n + D u_n \end{aligned}$$

A $n \times n$
B $n \times k$

We are given x_0, A, B, C, D. The sequence u_n $(n \geq 0)$ gives the inputs, the y_n are the outputs. We use generating functions to determine the relation between $(u_n)_{n \geq 0}$ and $(y_n)_{n \geq 0}$.

Set $X(z) = \sum^{T} x_n z^{-n}$, $U(z) = \sum_n^{T} u_n z^{-n}$, $Y(z) = \sum_n^{T} y_n z^{-n}$

Multiply the first equation in (1) by z^{-n-1} and sum:

$$X(z^{-1}) - x_0 = z^{-1} A X(z) + z^{-1} B U(z)$$

$$\rightarrow \quad (I - z^{-1}A) X(z^{-1}) = x_0 + z^{-1} B U(z)$$

$$X(t) = x_0 (I - z^{-1}A)^{-1} + z^{-1}(I - z^{-1}A)^{-1} B U(z)$$

Multiplying the second equation by z^n and summing yields:

$$Y(z) = C X(z) + D U(z)$$

$$= x_0 C (I - z^{-1}A)^{-1} + z^{-1} C (I - z^{-1}A)^{-1} B U(z) + D U(z)$$

Assume $x_0 = 0$. Then

$$Y(z) = (C (zI - A) B + D) U(z)$$

It's traditional to refer to

$$C(zI - A)^{-1} B + D$$

control theory
convolutional code

as the transfer function of our system.

Equivalence

We start with equivalence of 2×1 vectors over a ring R.

$$\begin{pmatrix} a & b \\ -y & x \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ 0 \end{pmatrix}$$

Here $ax+by$ lies in the ideal $\langle x, y \rangle$ generated by x & y. If R is a principal ideal domain, then $\langle x, y \rangle = d$, and we may choose a, b so that $ax+by = d$. Then, since $d|x$ & $d|y$

$$\begin{pmatrix} a & b \\ -y/d & x/d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}$$

As

$$\det\begin{pmatrix} a & b \\ -y/d & x/d \end{pmatrix} = 1$$

we see that the vectors

$$\begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} d \\ 0 \end{pmatrix}$$

are equivalent.

Lemma Assume R is a principal ideal domain. If $y = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in R^m$, then y is equivalent to $\gcd\{x_1, \dots, x_m\} e_1$

Lemma Assume R is a principal ideal domain. Then A is equivalent to a matrix of the form

$$\begin{bmatrix} \delta & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}.$$

Proof Using the previous lemma, A is equivalent to

$$\begin{bmatrix} d_1 & y_1 \\ 0 & \\ \vdots & B_1 \\ 0 & \end{bmatrix}$$

Each entry of this matrix lies in the ideal generated by the elements of A; assume this ideal is generated by δ.

Applying the row version of the previous lemma, we bring the previous matrix to the form

$$\left[\begin{array}{c|c} d_2 & 0 \cdots 0 \\ \hline z_2 & B_2 \end{array}\right]$$

where $d_2 | d_1$. If $d_2 = d_1$, then $z_1 = 0$.

So we continue; at each step either the $(1,1)$-entry decreases, or we're done. $\square$

We say an $m \times n$ matrix D is diagonal if $D_{i,j} = 0$ when $i \neq j$.

Existence of Smith normal form:

Theorem Any matrix over a principal ideal domains is equivalent to a diagonal matrix D such that $D_{i,i} \mid D_{i+1,i+1}$ for each i.

Proof Only the divisibility is in question.

We have

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow[\begin{bmatrix} 1 & 0 \\ x & 1 \end{bmatrix}_l]{} \begin{pmatrix} a & 0 \\ xa & b \end{pmatrix} \xrightarrow[\begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix}_r]{} \begin{pmatrix} a & 0 \\ \underbrace{xa+yb}_{\delta} & b \end{pmatrix}$$

$$\xrightarrow[\begin{pmatrix} 1 & -a/\delta \\ 0 & 1 \end{pmatrix}_l]{} \begin{pmatrix} 0 & ab/\delta \\ \delta & b \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & ab/\delta \\ \delta & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \delta & 0 \\ 0 & \frac{ab}{\delta} \end{pmatrix} \qquad \square$$

We are left with uniqueness.

Fitting ideals

Binet-Cauchy

Assume A & B are $k \times n$ matrices $k \leq n$.

If C is $k \times n$ and S is a k-subset of $\{1, \dots, n\}$, define $\rho_S(C)$ to be the determinant of the $k \times k$ matrix formed by the columns of C indexed by S.

Theorem

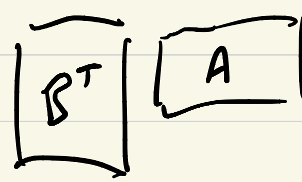
$$\det(AB^T) = \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| = k}} \rho_S(A)\, \rho_S(B)$$

The cases $k=1$ & $k=n$ are familiar.

Proof/ Recall that

$$\det(tI + B^TA) = t^{n-k}\det(tI + AB^T)$$

and consider the coefficient of t^{n-k}.
On the right, it's $\det(AB^T)$. On the left, we find the sum of the principal $k\times k$ minors of B^TA. Now if $|S|=k$ & $S = \{1,\dots,n\}$, the (S,S)-minor of B^TA is
$p_S(B)\,p(A)$.



Tao, via wikipedia

The Fitting ideal $F_k(A)$ of a matrix A is the ideal generated by the $(k\times k)$-minors of A.

It is defined over any commutative ring R.

Over a principal ideal domain, there are elements f_k of R such that $F_k = (f_k)$

Theorem If A and B are equivalent matrices, $F_k(A) = F_k(B)$ for all k.

Proof Assume P is invertible. By Binet-Cauchy, each $r \times r$ minor of PB lies in $F_r(B)$. Hence

$$F_r(PB) \leq F_r(B).$$

Since the entries of P^{-1} lie in R

$$F_r(B) = F_r(P^{-1}(PB)) \leq F_r(PB). \qquad \square$$

Finally, if D is a diagonal matrix & $D_{i,i} \mid D_{i+1,i+1}$ for each i, then $F_r(D) = D_{r,r}$

The terms $D_{i,i}$ are the invariant factors of A.

The invariant factors of $tI-A$ are referred to as its elementary divisors. The product $\prod D_{i,i}$ of the elementary divisors is $\phi(A,t)$; further $D_{n,n}$ is the minimal polynomial of A.

Division

We want to divide $F(t) = \sum_{k=0}^{r} F_k t^k$ by $tI - A$. We have

$$(tI-A)^{-1} = t^{-1} \sum_{i \geq 0} A^i t^{-i}$$

and hence the coefficient of t^{-1-j} in $F(t)(tI-A)^{-1}$ is

$$F_0 A^j + F_1 A^{j+1} + \dots + F_r A^{j+r} = (F_0 + F_1 A + \dots + F_r A^r) A^j$$

Therefore

$$F(t)(tI-A)^{-1} = \underbrace{Q(t)}_{\text{polynomial}} + F(A)(tI-A)^{-1}$$

and

$$\boxed{F(t) = Q(t)(I-tA) + F(A)}$$

$$(tI-A)\,\mathrm{adj}(tI-A) = \phi(t)\,I$$

$$\phi(t)I = Q(t)(tI-A) + \phi(A)$$

$$\left((tI-A) - Q(t)\right)(tI-A) = \phi(A)$$

$$\Rightarrow \phi(A) = 0$$

Cayley-Hamilton

Equivalence & Similarity

We redeem a debt by proving:

Theorem Let $\mathbb{F}$ be a field. Two $n \times n$ matrices over $\mathbb{F}$ are similar if & only if $tI-A$ and $tI-B$ are equivalent over $\mathbb{F}[t]$.

Proof You may prove that if A and B are similar, $tI-A$ and $tI-B$ are equivalent over $\mathbb{F}[t]$.

So assume we have invertible matrices $P(t)$ & $Q(t)$ such that

$$P(t)(tI-A) = (tI-B)Q(t)$$

Then we may write

$$P(t) = (tI - B)P_0(t) + P_1$$

$$Q(t) = Q_0(t)(tI - A) + Q_1$$

and therefore

$$0 = \big((tI - B)P_0(t) + P_1\big)(tI - A) - (tI - B)\big(Q_0(t)(tI - A) + Q_1\big)$$

$$= \underbrace{(tI - B)\big(P_0(t) - Q_0(t)\big)(tI - A)}_{\text{degree} \geq 2} + \underbrace{P_1(tI - A) - (tI - B)Q_1}_{\text{degree} \leq 1}$$

So $P_0(t) - Q_0(t)$ and

$$\boxed{P_1(tI - A) = (tI - B)Q_1}$$

We see that $P_1 = Q_1$ and similarity follows if we prove Q_1 is invertible. We may assume that

$$Q(t)^{-1} = R_0(t)(tI-B)+R_1$$

(Recall: $Q(t) = Q_0(t)(tI-A)+Q_1$)

We have

$$I = (R_0(tI-B)+R_1)(Q_0(tI-A)+Q_1)$$

$$= R_0(tI-B)R_1(tI-A) + R_1Q_0(tI-A) + R_0\overbrace{(tI-B)Q_1}^{=P_1(tI-A)} + R_1Q_1$$

$$= \underbrace{(R_0(tI-B)R_1 + R_1Q_0 + R_0P_1)(tI-A)}_{\text{degree} \geq 1} + \underbrace{R_1Q_1}_{\text{constant}}$$

Therefore $R_1 Q_1 = I$ & Q_1 (and P_1) are invertible.