Block LU-factorization of Hankel matrices

Henri, Nadia et moi

Besançon 2011

Outline

1. The origin
2. Block diagonalization of Hankel matrices
3. Hankel and Bezout matrices of two polynomials
4. Berlekamp–Massey
5. Bibliografy

The origin.

To understand the Proof of Frobenius' Theorem

Frobenius' Theorem

Let h be a regular Hankel matrix and $h_{\mathrm{k}_0} := 1$.

Let $h_{\mathrm{k}_1}, \ldots, h_{\mathrm{k}_q}$ be the non-zero principal leading minors of h $(1 \leq \mathrm{k}_1 < \cdots < \mathrm{k}_q = n)$.

Then the signature of h is given:

$$\mathrm{Sig}(h) = \sum_{i=1}^{q} \begin{cases} 0 & \text{if } k_i - k_{i-1} \text{ is even} \\ (-1)^{\frac{r_i - 1}{2}} \operatorname{sign}(h_{\mathrm{k}_i} h_{\mathrm{k}_{i-1}}) & \text{if } k_i - k_{i-1} \text{ is odd} \end{cases}$$

F. Gantmacher. *Theory of Matrices, Vol. 1,*, pag 343; Proof: 344–348.

Outline

1. The origin
2. Block diagonalization of Hankel matrices
3. Hankel and Bezout matrices of two polynomials
4. Berlekamp–Massey
5. Bibliografy

Block diagonalization of Hankel matrices–Preliminaries

A Hankel matrix h is

$$h = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_2 & \alpha_3 & & \iddots & \alpha_{n+1} \\ \alpha_3 & & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & & \vdots \\ \alpha_n & \alpha_{n+1} & \cdots & \cdots & \alpha_{2n-1} \end{pmatrix} = \mathrm{H}(\alpha_1, \ldots, \alpha_{2n-1}),$$

Block diagonalization of Hankel matrices–Preliminaries

Suppose that

$$
h = \mathrm{H}(\underbrace{0,\ldots,0}_{(r-1)},\alpha_r,\ldots,\alpha_{2n-1}) = \left(\begin{array}{ccc|ccc}
0 & \cdots & \alpha_r & \alpha_{r+1} & \cdots & \alpha_n \\
\vdots & \iddots & \vdots & \vdots & & \vdots \\
\alpha_r & \cdots & \alpha_{2r-1} & \vdots & & \vdots \\
\hline
\vdots & & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & & \vdots \\
\alpha_n & \cdots & \alpha_{n+r-1} & \alpha_{n+r} & \cdots & \alpha_{2n-1}
\end{array}\right)
$$

$$
= \left(\begin{array}{c|c} h_{11} & h_{12} \\ \hline h_{12}^t & h_{22} \end{array}\right),
$$

Block diagonalization of Hankel matrices–Preliminaries

Then $h = \mathrm{H}(\underbrace{0,\ldots,0}_{(r-1)},\alpha_r,\ldots,\alpha_{2n-1}) = \begin{pmatrix} h_{11} & h_{12} \\ h_{12}^t & h_{22} \end{pmatrix}$,

and consider

$$T = \begin{pmatrix} \alpha_r & \cdots & & \alpha_{2n-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & \alpha_r \end{pmatrix} = \mathrm{UpToep}(\alpha_r,\ldots,\alpha_{2n-1});$$

$$t = \mathrm{UpToep}(\alpha_r,\ldots,\alpha_{n+r-1}).$$

Block diagonalization of Hankel matrices

$h = \mathrm{H}(0, \ldots, 0, \alpha_r, \ldots, \alpha_{2n-1}),\ T = \mathrm{UpToep}(\alpha_r, \ldots, \alpha_{2n-1}),\ t = \mathrm{UpToep}(\alpha_r, \ldots, \alpha_{n+r-1})$

Block diagonalization of Hankel matrices

$h = \mathrm{H}(0,\ldots,0,\alpha_r,\ldots,\alpha_{2n-1})$, $T = \mathrm{UpToep}(\alpha_r,\ldots,\alpha_{2n-1})$, $t = \mathrm{UpToep}(\alpha_r,\ldots,\alpha_{n+r-1})$

Theorem. N. Ben–Atti et H. Lombardi

If $T^{-1} = \mathrm{UpToep}(\mu_1,\ldots,\mu_{2n-r})$, $t^{-1} = \mathrm{UpToep}(\mu_1,\ldots,\mu_n)$, then

$$h' = (t^{-1})^t\, h\, t^{-1} = \begin{pmatrix} h'_{11} & 0 \\ 0 & h'_{22} \end{pmatrix},$$

where

$$h'_{11} = J\, h_{1,1}^{-1} J = \mathrm{LowH}(\mu_1,\ldots,\mu_r) \qquad \text{and} \qquad h'_{22} = -\mathrm{H}(\mu_{r+2},\ldots,\mu_{2n-r}).$$

Block diagonalization of Hankel matrices

$h = \mathrm{H}(0, \ldots, 0, \alpha_r, \ldots, \alpha_{2n-1})$, $T = \mathrm{UpToep}(\alpha_r, \ldots, \alpha_{2n-1})$, $t = \mathrm{UpToep}(\alpha_r, \ldots, \alpha_{n+r-1})$

Theorem. N. Ben–Atti et H. Lombardi

If $T^{-1} = \mathrm{UpToep}(\mu_1, \ldots, \mu_{2n-r})$, $t^{-1} = \mathrm{UpToep}(\mu_1, \ldots, \mu_n)$, then

$$h' = (t^{-1})^t \, h \, t^{-1} = \begin{pmatrix} h'_{11} & 0 \\ 0 & h'_{22} \end{pmatrix},$$

where

$$h'_{11} = J \, h_{1,1}^{-1} J = \mathrm{LowH}(\mu_1, \ldots, \mu_r) \qquad \text{and} \qquad h'_{22} = -\mathrm{H}(\mu_{r+2}, \ldots, \mu_{2n-r}).$$

Corollary

There is an Upper Triangular matrix A and a block diagonal matrix, D, such that

$$A^t \, h \, A = \mathrm{Diag}(\mathrm{lh}_{11}, \ldots, \mathrm{lh}_{tt})$$

and every lh_{ii} is a lower Hankel matrix.

Block diagonalization–Euclidean Algorithm

Inversion of Toeplitz matrices: If

$$T = \mathrm{UpToep}(\alpha_1, \dots, \alpha_{2n-1}),\ \alpha_1 \neq 0,\, T^{-1} = \mathrm{UpToep}(\mu_1, \dots, \mu_{2n-1})$$

and we define

$$S(x) = \alpha_1 + \alpha_2 x + \ldots + \alpha_{2n-1} x^{2n-2}, Q(x) = \mu_1 + \mu_2 x + \ldots + \mu_{2n-1} x^{2n-1},$$

then $S(x)Q(x) = 1 \mod x^{2n-1}$

Block diagonalization–Euclidean Algorithm

Inversion of Toeplitz matrices: If

$$T = \mathrm{UpToep}(\alpha_1, \ldots, \alpha_{2n-1}),\ \alpha_1 \neq 0,\ T^{-1} = \mathrm{UpToep}(\mu_1, \ldots, \mu_{2n-1})$$

and we define

$$S(x) = \alpha_1 + \alpha_2 x + \ldots + \alpha_{2n-1} x^{2n-2}, Q(x) = \mu_1 + \mu_2 x + \ldots + \mu_{2n-1} x^{2n-1},$$

then $S(x)Q(x) = 1 \mod x^{2n-1}$

This leads the Algorithm for Computing the blocks lh_{ii}: Let

- $R_0 = x^{2n-1}; R_1 = \alpha_1 x^{2n-2} + \ldots + \alpha_{2n-1}$;
- $\{Q_1, \ldots, Q_{t-1}\}$: Quotients in the signed Euclidean algorithm for (R_0, R_1);

$$Q_i = c_{0,i} x^{d_1} + \ldots + c_{d_1,i}$$

- R_t: the last non zero remainder,

then

$$\boxed{\mathrm{lh}_{ii} = \mathrm{LowH}(c_{0,i}, \ldots, c_{d_i-1,i}), i \leq t-1; \mathrm{lh}_{tt} = \mathrm{LowH}(\text{ from } R_t)}$$

Proof of Frobenius' Theorem

$$A^t\, h\, A = D = \mathrm{Diag}(\mathrm{lh}_{11}, \mathrm{lh}_{22}, \ldots, \mathrm{lh}_{tt})$$

Proof of Frobenius' Theorem

$$A^t\, h\, A = D = \mathrm{Diag}(\mathrm{lh}_{11}, \mathrm{lh}_{22}, \ldots, \mathrm{lh}_{tt})$$

We have

$$\mathrm{Sig}(h) \overbrace{=}^{\text{Sylvester's law}} \mathrm{Sig}(D) = \mathrm{Sig}(\mathrm{Diag}(\mathrm{lh}_{11}, \ldots, \mathrm{lh}_{qq})) = \sum_{i=1}^{q} \mathrm{Sig}(\mathrm{lh}_{ii}). \qquad (1)$$

Since

$$\mathrm{Sig}(\mathrm{lh}_{ii}) = \begin{cases} 0 & \text{if } d_i \text{ is even,} \\ \begin{aligned} \mathrm{sign}(c_{0,i}) &= (-1)^{\frac{d_i-1}{2}}\,\mathrm{sign}(\det(\mathrm{lh}_{ii})) \\ &= (-1)^{\frac{d_i-1}{2}}\,\mathrm{sign}(\det(D_{k_i} D_{k_{i-1}})) \\ &= (-1)^{\frac{d_i-1}{2}}\,\mathrm{sign}(\det(h_{k_i} h_{k_{i-1}})) \end{aligned} & \begin{aligned} &\text{if } d_i \text{ is odd;} \\ &d_i = k_i - k_{i-1} \end{aligned} \end{cases} \qquad (2)$$

it follows the Frobenius' Theorem from (1) and (2)

$$\mathrm{Sig}(h) = \sum_{i=1}^{q} \begin{cases} 0 & \text{if } k_i - k_{i-1} \text{ is even} \\ (-1)^{\frac{r_i-1}{2}}\,\mathrm{sign}(h_{\mathrm{k}_i} h_{\mathrm{k}_{i-1}}) & \text{if } k_i - k_{i-1} \text{ is odd} \end{cases}$$

Literature

In [Gragg-Lindquist] (1983) and [Heinig-Rost] (1984), they assert that a Hankel matrix is LU-equivalent to a block diagonal matrix.

Theorem

There is an upper triangular matrix A_1 with entries equal to one on the main diagonal such that

$$A_1^t h A_1 = D_1$$

where D_1 is a block diagonal matrix and each block is a lower Hankel triangular matrix.

Such a block diagonalization can also be found for example in [Bini-Pan] and [Barel-Bultheel].

Outline

1 The origin

2 Block diagonalization of Hankel matrices

3 Hankel and Bezout matrices of two polynomials

4 Berlekamp–Massey

5 Bibliografy

Hankel and Bezout matrices of two polynomials

Let $u(x)$ and $v(x)$ be two coprime polynomials,

$$u(x) = \sum_{i=0}^{n} u_i x^i; \quad v(x) = \sum_{i=0}^{m} v_i x^i, \; n > m.$$

- The $n \times n$ Hankel matrix associated to $u(x)$ and $v(x)$ is defined as

$$\mathrm{H}(u, v) = \mathrm{H}\left(h_1, h_2, \ldots, h_{2n-1}\right),$$

where

$$R(x) = \frac{v(x)}{u(x)} = \sum_{i=1}^{\infty} h_i x^{-i}.$$

Hankel and Bezout matrices of two polynomials

Let $u(x)$ and $v(x)$ be two coprime polynomials,

$$u(x) = \sum_{i=0}^{n} u_i x^i; \quad v(x) = \sum_{i=0}^{m} v_i x^i, \; n > m.$$

- The $n \times n$ Hankel matrix associated to $u(x)$ and $v(x)$ is defined as

$$\mathrm{H}(u, v) = \mathrm{H}\left(h_1, h_2, \dots, h_{2n-1}\right),$$

  where

$$R(x) = \frac{v(x)}{u(x)} = \sum_{i=1}^{\infty} h_i x^{-i}.$$

- The Bezout Matrix associated to $u(x)$ and $v(x)$ is the symmetric matrix

$$\mathrm{Bez}(u, v) = \begin{pmatrix} c_{0,0} & \cdots & c_{0,n-1} \\ \vdots & & \vdots \\ c_{n-1,0} & \cdots & c_{n-1,n-1} \end{pmatrix},$$

  where the $c_{i,j}$ are defined by the Cayley expression:

$$\frac{u(x)v(y) - u(y)v(x)}{x - y} = \sum_{i,j=0}^{n-1} c_{i,j} x^i y^j.$$

Hankel and Bezout matrices of two polynomials

Properties

- $\text{degree}(\gcd(u, v)) = n - \mathsf{rk}(\text{Bez}(u, v)) = n - \mathsf{rk}(\text{H}(u, v))$.
- Both are symmetric.
- Both matrices provide subresultant polynomials.
- Both matrices represent the multiplication by v(x) modulo u(x).
- $\text{Bez}(u, v) = \text{Bez}(u, 1)\text{H}(u, v)\text{Bez}(u, 1)$, with

$$\text{Bez}(u, 1) = \text{UpperHk}(u_1, \dots, u_n).$$

  - $\text{J}\,\text{Bez}(u, v)\text{J} = \text{J}\,\text{Bez}(u, 1)\text{H}(u, v)\text{Bez}(u, 1)\,\text{J}$

$$\Downarrow \quad \text{H}(u, v) = A^{-t}\, D\, A^{-1}$$

    - $\text{J}\,\text{Bez}(u, v)\text{J} = \left[\text{J}\,\text{Bez}(u, 1)A^{-t}\right]\, D\, \left[A^{-1}\text{Bez}(u, 1)\,\text{J}\right]$

Block Diagonalization for $\mathrm{H}(u, v)$

If we apply our algorithm, in the first step

$$\boxed{(t^{-1})^t\, \mathrm{H}(u, v)\, t^{-1} = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1, 1)\,\mathrm{J} & 0 \\ 0 & \mathrm{H}(v, r_1) \end{pmatrix}}$$

with

$$u(x) = v(x)q_1(x) - r_1(x), \;\; t = \mathrm{UpToep}(h_{n-m}, \ldots, h_{2n-m}).$$

Block Diagonalization for $\mathrm{H}(u, v)$

If we apply our algorithm, in the first step

$$\boxed{(t^{-1})^t \,\mathrm{H}(u,v)\, t^{-1} = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1,1)\,\mathrm{J} & 0 \\ 0 & \mathrm{H}(v,r_1) \end{pmatrix}}$$

with

$$u(x) = v(x)q_1(x) - r_1(x),\ \ t = \mathrm{UpToep}(h_{n-m}, \ldots, h_{2n-m}).$$

In the last one,

$$\boxed{A^t \,\mathrm{H}(u,v)\, A = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1,1)\,\mathrm{J} & & & \\ & \mathrm{J}\,\mathrm{Bez}(q_2,1)\,\mathrm{J} & & \\ & & \ddots & \\ & & & \mathrm{H}(r_{t-1},r_t) \end{pmatrix}}$$

Block Diagonalization for $\mathrm{H}(u, v)$

If we apply our algorithm, in the first step

$$\boxed{(t^{-1})^t\,\mathrm{H}(u,v)\,t^{-1} = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1,1)\,\mathrm{J} & 0 \\ 0 & \mathrm{H}(v,r_1) \end{pmatrix}}$$

with

$$u(x) = v(x)q_1(x) - r_1(x),\ \ t = \mathrm{UpToep}(h_{n-m},\ldots,h_{2n-m}).$$

In the last one,

$$\boxed{A^t\,\mathrm{H}(u,v)\,A = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1,1)\,\mathrm{J} & & & \\ & \mathrm{J}\,\mathrm{Bez}(q_2,1)\,\mathrm{J} & & \\ & & \ddots & \\ & & & \mathrm{H}(r_{t-1},r_t) \end{pmatrix}}$$

Moreover, the rows of $\mathbf{A}^{-1}\,\mathrm{Bez}(\mathbf{u},\mathbf{1})\mathrm{J}$ are defined by the remainders.

Block Diagonalization for $\mathrm{Bez}(u, v)$

In [Bini-Gemignani] (1995),

$$\mathrm{JBez}(u,v)\mathrm{J} = \begin{pmatrix} (\mathrm{JBez}(u,v)\mathrm{J})_{m_i} & X^t \\ X & W \end{pmatrix}$$

$$= \begin{pmatrix} T & 0 \\ K & I \end{pmatrix} \begin{pmatrix} \mathrm{JBez}(u^{(i)}, v^{(i)})\mathrm{J} & 0 \\ 0 & \mathrm{JBez}(r_{i-1}, r_i)\mathrm{J} \end{pmatrix} \begin{pmatrix} T^t & K^t \\ 0 & I \end{pmatrix},$$

where

$H(u,v)_{m_i} = \mathrm{H}(u^{(i)}, v^{(i)})$, $\quad \mathrm{JBez}(r_{i-1}, r_i)\mathrm{J}$: **Schur complement** of $(\mathrm{JBez}(u,v)\mathrm{J})_{m_i}$,

Block Diagonalization for $\mathrm{Bez}(u, v)$

In [Bini-Gemignani] (1995),

$$\mathrm{JBez}(u,v)\mathrm{J} = \begin{pmatrix} (\mathrm{JBez}(u,v)\mathrm{J})_{m_i} & X^t \\ X & W \end{pmatrix}$$

$$= \begin{pmatrix} T & 0 \\ K & I \end{pmatrix} \begin{pmatrix} \mathrm{JBez}(u^{(i)}, v^{(i)})\mathrm{J} & 0 \\ 0 & \mathrm{JBez}(r_{i-1}, r_i)\mathrm{J} \end{pmatrix} \begin{pmatrix} T^t & K^t \\ 0 & I \end{pmatrix},$$

where

$$H(u,v)_{m_i} = \mathrm{H}(u^{(i)}, v^{(i)}), \quad \mathrm{JBez}(r_{i-1}, r_i)\mathrm{J} : \textbf{Schur complement} \text{ of } (\mathrm{JBez}(u,v)\mathrm{J})_{m_i},$$

If we sequentially apply the relation, then

$$\boxed{A_B^t\, \mathrm{J}\,\mathrm{Bez}(u,v)\,\mathrm{J}\, A_B = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1,1)\,\mathrm{J} & & & \\ & \mathrm{J}\,\mathrm{Bez}(q_2,1)\,\mathrm{J} & & \\ & & \ddots & \\ & & & \mathrm{J}\,\mathrm{Bez}(r_{t-1}, r_t)\,\mathrm{J} \end{pmatrix}}$$

Block Diagonalization for $\mathrm{Bez}(u, v)$

In [Bini-Gemignani] (1995),

$$\mathrm{JBez}(u,v)\mathrm{J} = \begin{pmatrix} (\mathrm{JBez}(u,v)\mathrm{J})_{m_i} & X^t \\ X & W \end{pmatrix}$$

$$= \begin{pmatrix} T & 0 \\ K & I \end{pmatrix} \begin{pmatrix} \mathrm{JBez}(u^{(i)}, v^{(i)})\mathrm{J} & 0 \\ 0 & \mathrm{JBez}(r_{i-1}, r_i)\mathrm{J} \end{pmatrix} \begin{pmatrix} T^t & K^t \\ 0 & I \end{pmatrix},$$

where

$H(u,v)_{m_i} = \mathrm{H}(u^{(i)}, v^{(i)})$, $\mathrm{JBez}(r_{i-1}, r_i)\mathrm{J}$: **Schur complement** of $(\mathrm{JBez}(u,v)\mathrm{J})_{m_i}$,

If we sequentially apply the relation, then

$$\boxed{A_B^t \,\mathrm{J}\,\mathrm{Bez}(u,v)\,\mathrm{J}\,A_B = \begin{pmatrix} \mathrm{J}\,\mathrm{Bez}(q_1,1)\,\mathrm{J} & & & \\ & \mathrm{J}\,\mathrm{Bez}(q_2,1)\,\mathrm{J} & & \\ & & \ddots & \\ & & & \mathrm{J}\,\mathrm{Bez}(r_{t-1}, r_t)\,\mathrm{J} \end{pmatrix}}$$

Moreover, the rows of $\mathbf{A_B^{-1}}$ are defined by the remainders.

Outline

1. The origin

2. Block diagonalization of Hankel matrices

3. Hankel and Bezout matrices of two polynomials

4. Berlekamp–Massey

5. Bibliografy

Berlekamp–Massey

Classical

We have: The first $2n$ coefficients of a linearly recurrent sequence, $[a_0, a_1, \ldots, a_{2n-1}]$. The minimal polynomial has degree bound n.

We want to compute: The minimal polynomial of the sequence P(x).

INPUT:

$$R_0 := x^{2n}, R_1 := a_0 + a_1\,x + \ldots + a_{2n-1}\,x^{2n-1};\, V_0 = 0;\, V_1 = 1;$$

GOAL: V such that $\mathrm{degree}(V\,R_1) < n$

ALGORITHM:
while $n \leq \mathrm{degree}(R_1)$ do

$(Q, R) :=$ quotient and remainder of R_0 divided by R_1 ;
$V := V_0 - Q\,V_1$;
$V_0 := V_1$; $V_1 := V$; $R_0 := R_1$; $R_1 := R$;

end while;

$d := \text{máx}(\mathrm{degree}(V_1), 1 + \mathrm{degree}(R_1))$; $\mathbf{P} := \mathbf{x^d\,V_1(1/x)}$;

Berlekamp–Massey

Modified

We have: The first $2n$ coefficients of a linearly recurrent sequence, $[a_0, a_1, \ldots, a_{2n-1}]$. The minimal polynomial has degree bound n.

We want to compute: The minimal polynomial of the sequence P(x).

INPUT:

$$R_0 := x^{2n}, R_1 := a_0 x^{2n-1} + a_1\, x^{2n-2} + \ldots + a_{2n-1}\,;\, V_0 = 0;\, V_1 = 1;$$

GOAL: V such that $\mathrm{degree}(V\, R_1) < n$

ALGORITHM:
while $n \leq \mathrm{degree}(R_1)$ do

$(Q, R) :=$ quotient and remainder of R_0 divided by R_1 ;
$V := V_0 - Q\, V_1$;
$V_0 := V_1$; $V_1 := V$; $R_0 := R_1$; $R_1 := R$;

end while;

$\mathbf{P} := \mathbf{V_1}/\mathbf{lc}(\mathbf{V_1})$;

Outline

1. The origin
2. Block diagonalization of Hankel matrices
3. Hankel and Bezout matrices of two polynomials
4. Berlekamp–Massey
5. Bibliografy

D. Bini and V. Pan.
Polynomial and Matrix Computations. Birkhäuser (1994).

A. Bultheel and M. Van Barel.
Linear Algebra, rational approximation and orthogonal polynomials.
Studies in Computational Mathematics 6, Elsevier (1997).

G. Heinig and K. Rost.
Algebraic Methods for Toeplitz-like Matrices and Operators. Birkhuser Verlag, Basel. (1984).

V. Shoup,
A Computational Introduction to Number Theory and Algebra,
Cambridge University Press (2005).

N. Ben Atti
Calcul rapide sur les matrices structurées : les matrices de Hankel.
Thèse, Université de Franche–Comté, (2008).

D. Bini and L. Gemignani.
Fast parallel computations of the polynomial remainder sequence via Bezout and Hankel matrix.
SIAM J. Comput. 24, 63–77 (1995).

W.B. Gragg and A. Lindquist.
On the partial realization problem. Linear Algebra Appl. 50, 277–319 (1983).

Merci Henri

Merci Henri

Merci de votre attention