MODIFIED MOMENTS FOR INDEFINITE WEIGHT FUNCTIONS

GENE H. GOLUB
MARTIN H. GUTKNECHT

AIM: Compute the recurrence coefficients of a set of orthogonal polynomials (belonging to some unknown measure λ).
Assume that for some sequence $\{\tau_n\}$ of polynomials the

- modified moments $\nu_n := \int \tau_n(z)\, d\lambda(z)$

or the

- Gramian $[\nu_{m,n}]_{m,n=0}^{\infty}$, $\nu_{m,n} := \int \tau_m(z) \tau_n(z)\, d\lambda(z)$

are known.

[Even in the classical case $\tau_n(z) = z^n$ of given ordinary moments $\mu_n := \int z^n \, d\lambda(z)$ some of our results seem to be new.]

APPLICS.:
- Parameter adaptation of iterative methods for linear systems [GOLUB/KENT]
- Quadrature
- EVALS ?

PREV. WORK: CHEBYSHEV 1859, SACK/DONOVAN '72, WHEELER '74, GAUTSCHI ['70, '78/84, '82, '86, ...]

1. Theory for the generic case

λ complex measure defined on subset of $\mathbb{C}$
(The subset may be $\mathbb{R}$, and λ may be real; "classical case": λ positive, on subset of $\mathbb{R}$.)

Assumption: All moments

(1.1) $$\mu_n := \int z^n \, d\lambda(z) \qquad (n \in \mathbb{N})$$

and all formal orthogonal polynomials (of the first kind), or, briefly, FOP1s, $\pi_n \in \mathcal{P}_n$ (of exact degree n, normalized to be monic) exist:

(1.2) $$\boxed{\int \pi_m(z)\pi_n(z)\, d\lambda(z) = \begin{cases} 0 & \text{if } m \neq n \\ \sigma_{n,n} \neq 0 & \text{if } m = n \end{cases}}$$

Define linear functional φ on $\mathcal{P} := \bigcup_{n=0}^{\infty} \mathcal{P}_n$ by

(1.1') $$\varphi(z^n) := \mu_n$$

Then (1.2) becomes

(1.2') $$\boxed{\varphi(\pi_m \pi_n) = \begin{cases} 0 & \text{if } m \neq n \\ \sigma_{n,n} \neq 0 & \text{if } m = n \end{cases}}$$

N.B.: $\varphi : (p,q) \mapsto \varphi(pq)$ is not an inner product.

[For an inner product one would need a positive measure and

$$\varphi(p,q) = \int \overline{p(z)}\, q(z)\, d\lambda(z)$$
]

(1.2') implies

(1.3) $$\varphi(p\pi_n) = 0 \qquad \forall p \in \mathcal{P}_{n-1}$$

The assumption implies three-term recurrence for $\{\pi_n\}$:

(1.4) $$\pi_{n+1}(z) = (z - \alpha_n)\pi_n(z) - \beta_n \pi_{n-1}(z), \quad n = 0, 1, \ldots$$

(with $\pi_{-1}(z) :\equiv 0$)

Consider second set $\{\tau_n\}$ of monic polynomials. They satisfy recurrence

(1.6) $$\tau_{n+1}(z) = z\,\tau_n(z) - \sum_{k=0}^{n} \tau_{k,n}\,\tau_k(z), \quad n = 0, 1, \ldots$$

We are most interested in the case where they too satisfy a three-term recurrence, e.g., where τ_n is a suitably shifted and normalized Chebyshev polynomial.

Modified moments:

(1.7) $$\nu_m := \varphi(\tau_n) = \int \tau_m(z)\, d\lambda(z) \qquad (m \in \mathbb{N})$$

More generally, following Chebyshev [1859], consider case $\tau_m(z) = z^m$

(1.8) $$\sigma_{m,n} := \varphi(\tau_m \pi_n) = \int \tau_m(z)\pi_n(z)\, d\lambda$$

Then

(1.9/10) $$\sigma_{m,n} = \begin{cases} 0 & \text{if } m < n \\ \varphi(\pi_n^2) \neq 0 & \text{if } m = n \\ \nu_m & \text{if } n = 0 \end{cases}$$ nonsingular lower triangular

$[\sigma_{m,n}]_{m,n=0}^{\infty}$ is the transposed of Gautschi's matrix Σ. Here, the modified moments ν_m are in the first column.

Matrix notation:

(1.11) $$\underset{\sim}{p} := [\pi_0, \pi_1, \pi_2, \dots], \qquad \underset{\sim}{t} := [\tau_0, \tau_1, \tau_2, \dots]$$

(1.12) $$G := \begin{bmatrix} \alpha_0 & \beta_1 & & \\ 1 & \alpha_1 & \beta_2 & \\ & 1 & \alpha_2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}, \qquad T := \begin{bmatrix} \tau_{00} & \tau_{01} & \tau_{02} & \dots \\ 1 & \tau_{11} & \tau_{12} & \\ & 1 & \tau_{22} & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

(1.13) $$S := \begin{bmatrix} \sigma_{00} & & & \\ \sigma_{10} & \sigma_{11} & & \\ \sigma_{20} & \sigma_{21} & \sigma_{22} & \\ \vdots & & & \ddots \end{bmatrix}, \qquad D := \begin{bmatrix} \sigma_{00} & & & \\ & \sigma_{11} & & \\ & & \sigma_{22} & \\ & & & \ddots \end{bmatrix}$$

The recurrence relations turn into

(1.14) $$\boxed{z\,\underset{\sim}{p}(z) = \underset{\sim}{p}(z)\,G}, \qquad \boxed{z\,\underset{\sim}{t}(z) = \underset{\sim}{t}(z)\,T}$$

and (1.2), (1.8) become

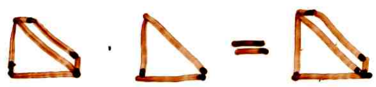
(1.15) $$\boxed{\varphi(\underset{\sim}{p}^T \underset{\sim}{p}) = D}, \qquad \boxed{\varphi(\underset{\sim}{t}^T \underset{\sim}{p}) = S}$$

element by element application of φ to infinite rank-one matrices

There follows

$$SG = \varphi(\underset{\sim}{t}^T \underset{\sim}{p})G = \varphi(\underset{\sim}{t}^T \underset{\sim}{p} G) = \varphi(\underset{\sim}{t}^T z \underset{\sim}{p}) =$$
$$= \varphi(T^T \underset{\sim}{t}^T \underset{\sim}{p}) = T^T \varphi(\underset{\sim}{t}^T \underset{\sim}{p}) = T^T S,$$

i.e.,

(1.16)  $$\boxed{SG = T^T S}$$ 

More matrix notation:

$$\underset{\sim}{z}(z) := [1, z, z^2, \ldots]$$

(1.17) $$M := [\mu_{m+n}]_{m,n=0}^{\infty} = \varphi(\underset{\sim}{z}^T \underset{\sim}{z})$$ moment matrix (Hankel)

(1.18) $$N := [\varphi(\tau_m \tau_n)]_{m,n=0}^{\infty} = \varphi(\underset{\sim}{t}^T \underset{\sim}{t})$$ "formal" Gramian of $\{\tau_m\}$

(1.19) $$\underset{\sim}{z}(z) = \underset{\sim}{p}(z)\, Z$$

(1.20) $$\underset{\sim}{t}(z) = \underset{\sim}{p}(z)\, R$$

Z and R are unit upper triangular and contain in their nth columns the coefficients of z^n and τ_n, resp., in terms of the π_m $(m \le n)$

Inserting (1.19) into (1.17) yields

(1.21) $$\boxed{M = Z^T D Z}$$ LDU decomposition of M

By (1.15) and (1.20),

(1.22) $$\underline{S = R^T D}$$ S and R^T are the same up to column scaling

Inserting (1.20) and (1.22) into (1.18) yields in analogy to (1.21):

(1.23) $$\boxed{N = R^T D R = S R}$$

LDU LU decomposition of N

Our assumptions imply that these decompositions exist.

If M and N are positive definite, these are the square-root-free <u>Cholesky decompositions</u> of M and N.

If π_n are normalized by $\varphi(\pi_n^2) = 1$, so that $D = I$, $S = R$, and one has Cholesky decomp. in the positive case.

2. Algorithms for the generic case

(i) Equating the (m,n) element of $SG = T^T S$ yields:

(2.1) $$\boxed{\sigma_{m,n+1} + \alpha_n \sigma_{m,n} + \beta_n \sigma_{m,n-1} = \sigma_{m+1,n} + \sum_{k=n}^{m} \tau_{k,m} \sigma_{k,n}}$$

(with $\sigma_{m,-1} := 0$) $(m, n \in \mathbb{N})$

$m < n-1$: $0 = 0$

$m = n-1$: $\beta_n \sigma_{n-1,n-1} = \sigma_{n,n}$

(2.4a) $$\Longrightarrow \quad \boxed{\beta_n = \frac{\sigma_{n,n}}{\sigma_{n-1,n-1}}}$$

$m = n$: $\alpha_n \sigma_{n,n} + \beta_n \sigma_{n,n-1} = \sigma_{n+1,n} + \tau_{n,n} \sigma_{n,n}$

(2.4b) $$\Longrightarrow \quad \boxed{\alpha_n = \tau_{n,n} + \frac{\sigma_{n+1,n}}{\sigma_{n,n}} - \frac{\sigma_{n,n-1}}{\sigma_{n-1,n-1}}}$$

Hence:

$$\boxed{\left.\begin{matrix} \sigma_{00}, \sigma_{1,1}, \dots, \sigma_{\bar{m},\bar{m}} \\ \sigma_{1,0}, \sigma_{2,0}, \dots, \sigma_{\bar{m}+1,\bar{m}} \end{matrix}\right\} \Longrightarrow \left\{\begin{matrix} \beta_0, \beta_1, \dots, \beta_{\bar{m}} \\ \alpha_0, \alpha_1, \dots, \alpha_{\bar{m}} \end{matrix}\right.}$$

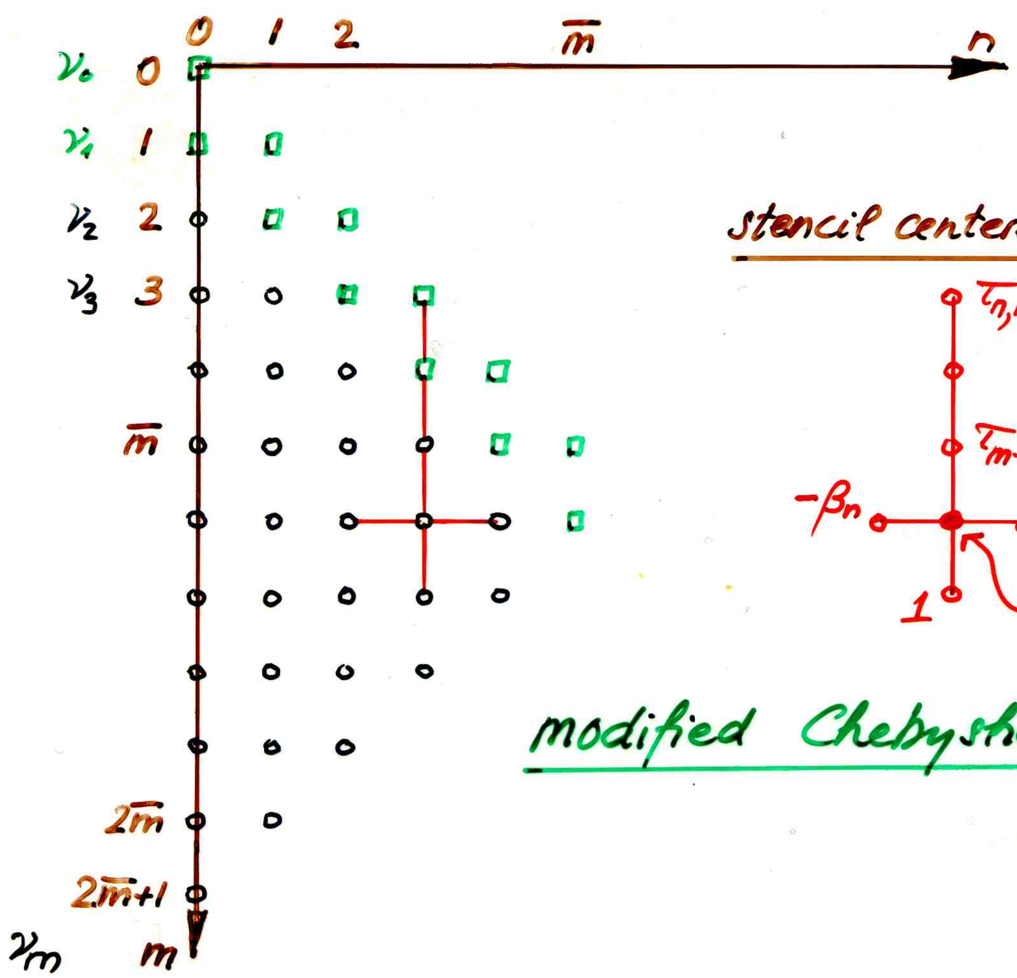
$\beta_0 := \nu_0 = \sigma_{0,0}$

(ii) Simultaneously, one can build up $S = [\sigma_{m,n}]$ from left to right (to its diagonal) using (1.25):

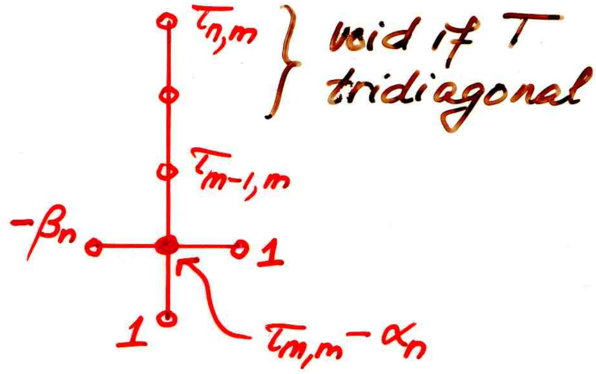
$$\boxed{\begin{array}{l} \sigma_{m,0} := \nu_m, \quad m = 0, 1, \dots, 2\bar{m}+1 \\ \text{for } n = 0, 1, \dots, \bar{m} \left\{ \begin{array}{l} \beta_n, \alpha_n \text{ acc. to (2.4a,b)} \; [n=0: \beta_0 = \nu_0] \\ \sigma_{m,n+1} \text{ acc. to (2.1)}, \; m = n+1, n+2, \dots, 2\bar{m}+1-n \end{array} \right. \end{array}}$$

Hence, $$\boxed{\nu_0, \nu_1, \dots, \nu_{2\bar{m}+1} \Longrightarrow \{\sigma_{m,n}\}_{n \le m = 0}^{\bar{m}}, \; \begin{matrix} \beta_0, \beta_1, \dots, \beta_{\bar{m}} \\ \alpha_0, \alpha_1, \dots, \alpha_{\bar{m}} \end{matrix}}$$

Modified Chebyshev algorithm [Sack/Donovan '72, Wheeler '74, Gautschi]



stencil centered at (m,n):



modified Chebyshev algorithm

(ii') Conversely:

$$\text{for } n = 0,1,\ldots,\bar m \quad \left\{ \begin{array}{l} \sigma_{n,n} := \sigma_{n-1,n-1}\,\beta_0 \qquad [n=0:\ \sigma_{n,n}=\beta_0:=\nu_0] \\ \sigma_{n+1,n} := (\alpha_n - \tau_{n,n})\,\sigma_{n,n} + \beta_n\,\sigma_{n,n-1} \end{array} \right.$$

$$\text{for } m = 0,1,\ldots,2\bar m \quad \left\{ \ \sigma_{m+1,n} \text{ acc. to (2.1)}, \quad n = 0,1,\ldots,\min\{m-1, 2\bar m - m\} \right.$$

Hence:

$$\left.\begin{array}{l} \beta_0, \beta_1, \ldots, \beta_{\bar m} \\ \alpha_0, \alpha_1, \ldots, \alpha_{\bar m} \end{array}\right\} \Longrightarrow \{\sigma_{m,n}\}_{n \le m = 0}^{\bar m}, \quad \nu_0 = \sigma_{0,0}, \ldots, \nu_{2\bar m} = \sigma_{2\bar m,0}$$

Inverse modified Chebyshev algorithm

(iii) LU decomposition of N: $N = SR$ [Gautschi '70]

$$\left.\begin{array}{l}(\bar m+2)\times(\bar m+1)\\ \text{principal}\\ \text{submatrix}\\ \text{of } N=[\nu_{m,n}]\end{array}\right\} \xRightarrow{\text{LU dec.}} \left\{\begin{array}{l}(\bar m+2)\times(\bar m+1)\\ \text{principal}\\ \text{submatrix}\\ \text{of } S\end{array}\right\} \begin{array}{l}\overset{(2.4)}{\nearrow} \left\{\begin{array}{l}\beta_0,\beta_1,\dots,\beta_{\bar m}\\ \alpha_0,\alpha_1,\dots,\alpha_{\bar m}\end{array}\right\}\\ \underset{(2.1)}{\searrow} \left\{\nu_{\bar m+1},\dots,\nu_{2\bar m+1}\right\}\end{array}$$

cf. Fig. below

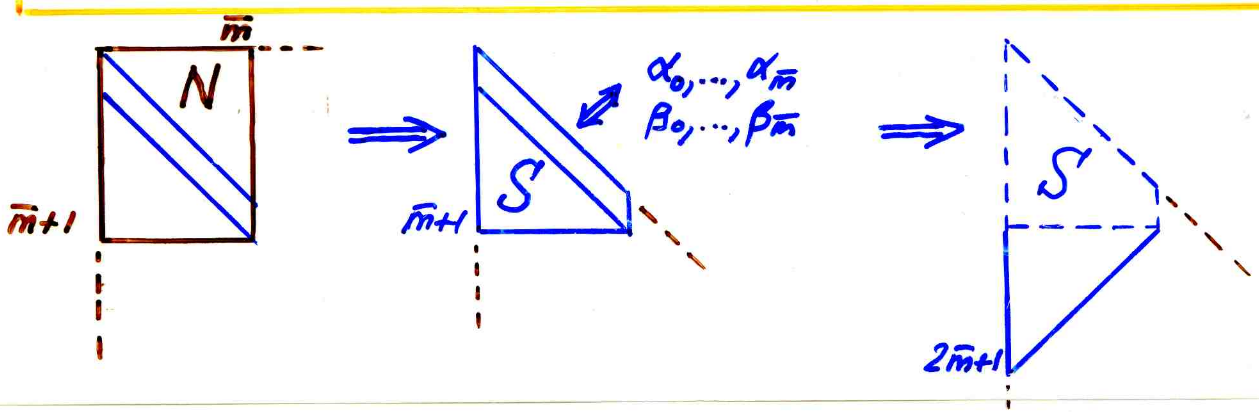
(iv) Recursive generation of S from diagonal + subdiagonal of N

Recall $S = R^T D$, $\quad N = SR = S D^{-1} S^T$

$$\text{for } m = 0,1,\dots,\bar m \quad \left\{\begin{array}{ll}\sigma_{m,m} := \nu_{m,m} - \sum_{k=0}^{m-1} \sigma_{m,k}^2\, \sigma_{k,k}^{-1} & \text{Chol.}\\ \sigma_{m+1,n} \text{ acc. to } (2.1),\ n=0,1,\dots,m-1 & \text{rec.}\\ \sigma_{m+1,m} := \nu_{m+1,m} - \sum_{k=0}^{m-1} \sigma_{m,k}\,\sigma_{m+1,k}\,\sigma_{k,k}^{-1} & \text{Chol.}\\ \beta_m, \alpha_m \text{ acc. to } (2.4a,b) & \text{expl.}\end{array}\right.$$

Subsequently, one can generate the lower half of the triangle of S in the figure and, hence, $\nu_{\bar m+2},\dots,\nu_{2\bar m+1}$, according to (2.1) [As in alg. (iii).]

$$\left.\begin{array}{l}\nu_{0,0},\nu_{11},\dots,\nu_{\bar m,\bar m}\\ \nu_{1,0},\nu_{21},\dots,\nu_{\bar m+1,\bar m}\end{array}\right\} \Rightarrow \{\sigma_{m,n}\}_{\substack{m\le \bar m+1\\ n\le \bar m}},\ \{\alpha_m,\beta_m\}_{m\le\bar m} \overset{(2.1)}{\Longrightarrow} \{\nu_m\}_{\bar m+1}^{2\bar m+1}$$



3. Matrix relations for the general case

φ : arbitrary complex linear functional defined on $\mathbb{P}$

$\mu_n := \varphi(z^n) \quad (n \in \mathbb{N})$

DEF. A monic polynomial π_n of exact degree n is called a regular formal orthogonal polynomial [regular FOP1] for φ if

(i) $\varphi(p\,\pi_n) = 0 \quad \forall p \in \mathbb{P}_{n-1}$

(ii) π_n is uniquely determined by (i)

It is called singular FOP1 if only (i) holds.
It is called deficient FOP1 if no regular or singular FOP1 of degree n exists and there holds

$$\varphi(p\,\pi_n) = 0 \quad \forall p \in \mathbb{P}_m$$

with the maximum possible value of $m (< n-1)$.

Let $0 = n_0 < n_1 < n_2 < \ldots < n_J \quad (J \leq \infty)$ be the finite or infinite sequence of indices for which a regular FOP1 π_{n_j} exists. Let $h_j := n_{j+1} - n_j$.

From Padé theory it is known [Gragg '72, Draux '83]:

Any singular or deficient FOP1 with $n_j < n < n_{j+1}$ has the form

$$\tilde{\pi}_n(z) = \omega(z)\,\pi_{n_j}(z),$$

where ω is an arbitrary (monic) polyn. of exact degree $n - n_j$

$\tilde{\pi}_n$ is singular iff $n_j < n \leq n_j + \left\lfloor \frac{h_j - 1}{2} \right\rfloor$

(3.1) May choose $\pi_n(z) := z^{n-n_j}\,\pi_{n_j}(z) \qquad (n_j \leq n < n_{j+1})$

THM [[Gragg '74], Draux '83, G.] ("Orthogonality")
Let π_n be chosen as above, and set $\underset{\sim}{p} := [\pi_0, \pi_1, \dots]$.
Then

(3.3)
$$\boxed{\varphi(\underset{\sim}{p}^T \underset{\sim}{p}) = D := \text{block diag}[D_0, D_1, D_2, \dots]}$$

with blocks D_j of size $h_j \times h_j$ which are right lower triangular Hankel matrices:

(3.4)
$$D_j := \begin{bmatrix} & & & & \delta_j \\ & & & \delta_j & \vdots \\ & & \cdot^{\cdot^{\cdot}} & & * \\ & \delta_j & \cdots & & \\ \delta_j & * & \cdots & * & * \end{bmatrix}, \quad \delta_j := \varphi(z^{h_j - 1}\pi_{n_j}^2).$$

If $J < \infty$, D_J is the infinite zero matrix.

If again $\underset{\sim}{z}(z) := [1, z, z^2, \dots]$, $\underset{\sim}{z}(z) = \underset{\sim}{p}(z) Z$, and $M := \varphi(\underset{\sim}{z}^T \underset{\sim}{z})$, then

(3.5)
$$\boxed{M = Z^T D Z}$$

is a block LDU decomposition with unit upper triangular Z.

THM [[Draux '83], G.]
For any equivalent set $\{\tilde{\pi}_n\}$ of FOP1s for φ there still holds (3.3) with symmetric (but not Hankel) blocks of the form (3.4). (3.5) too persists.

There is a particular set for which D_j is anti-diagonal ($j = 0, 1, \dots, J-1$).

THM. [(Gragg '74), Draux '83, G.] ("Recurrence")

There holds $\boxed{z\,\underset{\sim}{p}(z) = \underset{\sim}{p}(z)\,G}$, where G is an infinite block tridiagonal unit upper Hessenberg matrix,

(3.6)
$$G := \begin{bmatrix} A_0 & B_1 & & \\ C_0 & A_1 & B_2 & \\ & C_1 & A_2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

in which, for $0 \le i < J$, A_i is a companion matrix of order h_i,

(3.7)
$$A_i := \left[\begin{array}{cccc|c} 0 & & & & \\ 1 & \ddots & & & \\ & \ddots & \ddots & & a_i \\ & & \ddots & 0 & \\ & & & 1 & \end{array}\right], \qquad a_i := \begin{bmatrix} \alpha_{0,i} \\ \alpha_{1,i} \\ \vdots \\ \alpha_{h_i-1,i} \end{bmatrix},$$

while B_i and C_{i-1} have the form

(3.8)
$$B_i := \begin{bmatrix} 0 & \cdots & 0 & \beta_i \\ \vdots & & \vdots & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}, \; (\beta_i \neq 0) \qquad C_{i-1} := \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}.$$

If $J < \infty$, A_J is the infinite forward shift matrix and $B_J = C_{J-1}^T$ is the $h_{J-1} \times \infty$ zero matrix

As before, consider second set $\{\tau_n\}$ of monic polynomials, and let

$$\underset{\sim}{t} := [\tau_0, \tau_1, \tau_2, \ldots], \qquad z\,\underset{\sim}{t}(z) = \underset{\sim}{t}(z)\, T \quad \text{– unit upper Hessenberg}$$

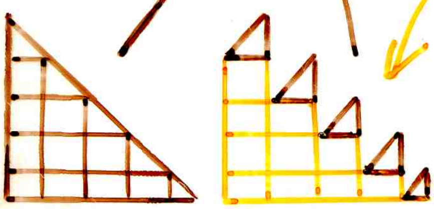
$$\underset{\sim}{p} := [\pi_0, \pi_1, \pi_2, \ldots], \qquad \underset{\sim}{t}(z) = \underset{\sim}{p}(z)\, R \quad \text{– unit upper triangular}$$

Again, there follow the relations

$$\boxed{S = R^T D\,, \quad N = R^T D R = SR\,, \quad SG = T^T S} \tag{3.9}$$

We partition also S, R, N, T into blocks of size $h_i \times h_j$ (as in G and D).

$S = R^T D \Rightarrow$ <u>S is lower block triangular with diagonal blocks of the form</u>



$$S_{jj} = \begin{bmatrix} & & & & \delta_j \\ & & & \delta_j & * \\ & & \cdot^{\cdot^{\cdot}} & & \vdots \\ & \delta_j & & & * \\ \delta_j & * & \cdots & * & * \end{bmatrix} \qquad \left(\begin{array}{c}\text{same form} \\ \text{as } D_j\end{array}\right)$$

T has the form

$$T = \begin{bmatrix} T_{00} & T_{01} & T_{02} & \cdots \\ C_0 & T_{11} & T_{12} & \cdots \\ & C_1 & T_{22} & \\ & & \ddots & \ddots \end{bmatrix}$$

with C_j as in G.

4. Algorithms for the general case

(i). In view of $SG = T^T S$ there holds now in analogy to (2.1) for the blocks

(4.1)
$$\boxed{\begin{aligned} &S_{i,j+1} C_j + S_{i,j} A_j + S_{i,j-1} B_j = \\ &\quad = C_i^T S_{i+1,j} + T_{i,i}^T S_{i,j} + \sum_{\ell=j}^{i-1} T_{\ell,i}^T S_{\ell,j} \end{aligned}}$$

$(i,j \geq 0)$

$i = j-1$:

(4.2)
$$S_{j-1,j-1} B_j = C_{j-1}^T S_{j,j}$$

$i = j$

(4.3)
$$S_{j,j} A_j + S_{j,j-1} B_j = C_j S_{j+1,j} + T_{j,j} S_{j,j}$$

Multiplication from the left with $S_{j-1,j-1}^{-1}$ and $S_{j,j}^{-1}$ would yield formulas for B_j and A_j.
However, since we know the structure of B_j and A_j, these formulas contain a lot of redundant information.
Let

$$S_{j-1,j} = \left[\begin{array}{c|cccc} & * & \cdots & * & \sigma_{j-1}' \\ s_j' & \vdots & & & * \\ & \vdots & & & \vdots \\ & * & \cdots & * & * \end{array}\right], \qquad S_{j,j} = \left[\begin{array}{cccc|c} & & & \delta_j & \\ & & \cdot & * & \\ & \delta_j & \cdot & * & s_j \\ \delta_j & * & \cdots & * & \end{array}\right],$$

then one obtains

(4.4), (4.6)
$$\boxed{\begin{aligned} \beta_j &= \delta_j / \delta_{j-1} \\ a_j &= S_{j,j}^{-1} \left[T_{j,j}^T s_j + e_{k_j} \sigma_j' - s_j' \beta_j \right] \end{aligned}}$$

$$e_{k_j} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Hence, $\boxed{\{S_{j,j}, s_j', \sigma_j'\}_{j=0}^{\bar{m}} \Longrightarrow \{a_j, \beta_j\}_{j=0}^{\bar{m}}}$

(i') In addition, one can still use the relation $SG = T^T S$ for an element by element comparison (in contrast to the block by block comparison). There follows:

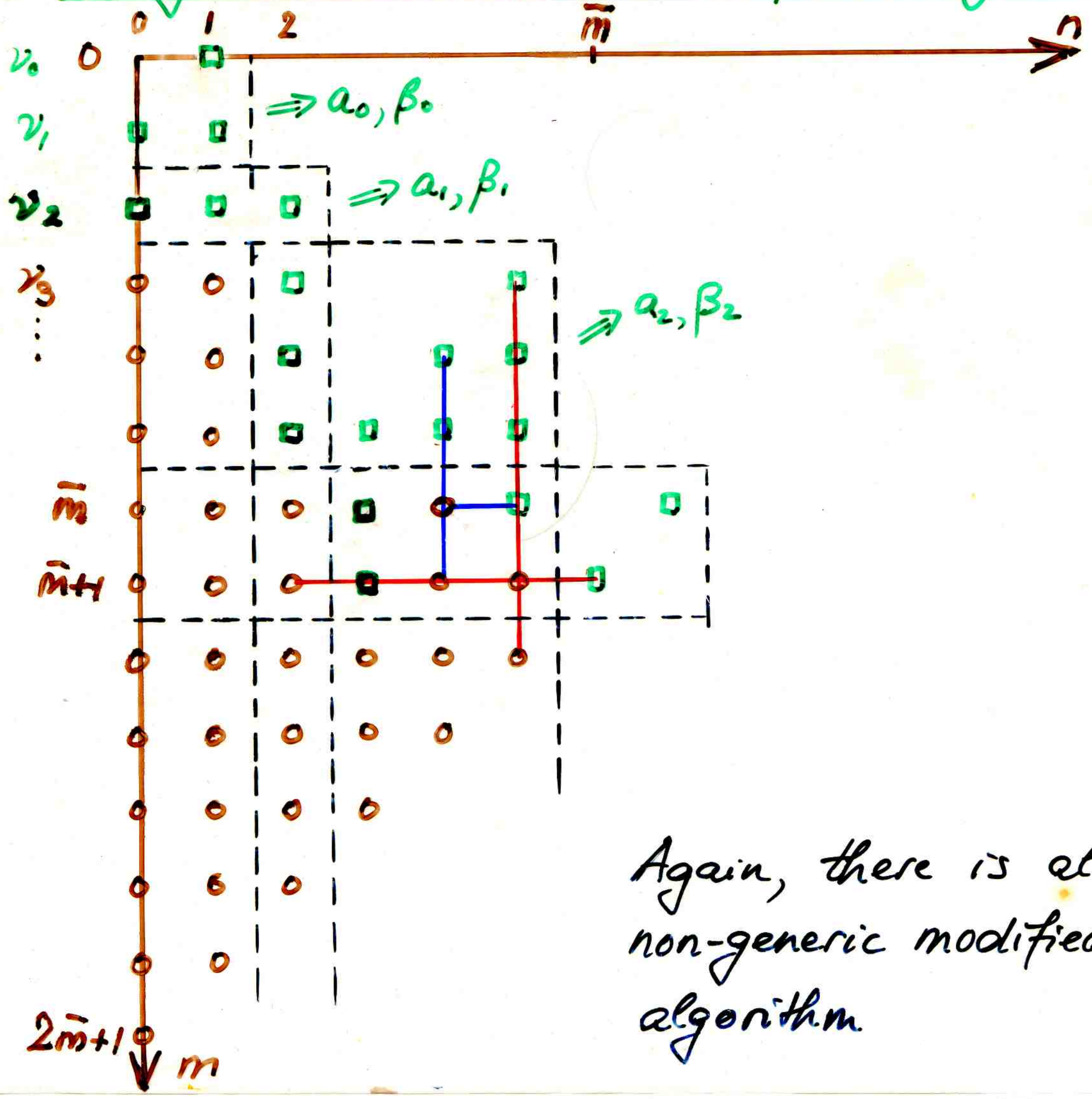
If $n_j \le n = n_j + k < n_{j+1} - 1$:

$$\sigma_{m,n+1} = \sigma_{m+1,n} + \sum_{\ell=0}^{m+k-n_{j+1}+1} \tau_{m-\ell,m}\, \sigma_{m-\ell,n} \tag{4.8a}$$

If $n = n_{j+1} - 1$:

$$\sigma_{m,n_{j+1}} + \sum_{\ell=1}^{h_j} \sigma_{m,n_{j+1}-\ell}\, \alpha_{h_j-\ell,j} + \sigma_{m,n_j-1}\, \beta_j = \sigma_{m+1,n_{j+1}-1} + \sum_{k=n_j}^{m} \tau_{k,m}\, \sigma_{k,n_{j+1}-1} \tag{4.8b}$$

(ii) Non-generic modified Chebyshev algorithm



Again, there is also an inverse non-generic modified Chebyshev algorithm.

MOTIVATION: Iterative solution of nonsymmetric system of linear equations

Polynomial acceleration method ($\equiv$ semiiterative method $\equiv$ gradient method) for solving $Az = b$ generates approximations z_n of the solution z s.t.

- residuals $r_n := b - Az_n$
- errors $e_n := z_n - z$

satisfy

$$r_n = \pi_n(A) r_0 , \qquad e_n = \pi_n(A) e_0 ,$$

where

$$\pi_n \in \mathcal{P}_n , \qquad \pi_n(0) = 1 .$$

- Need $\{\pi_n\}$ such that $\sup_{\lambda \in \sigma(A)} |\pi_n(\lambda)| \to 0$
- For efficiency, need $\{\pi_n\}$ s.t. 3-term (or k-term) recurrence relation holds ($\Rightarrow z_n$ satisfy such a rec.).
- If EVALs were known, one could choose $\{\pi_n\}$ s.t. π_ν = minimal polynomial of A ($\Rightarrow r_\nu = 0$)
- If domain Ω containing the EVALs is known, there are ways to construct a method appropriate for all matrices A with $\sigma(A) \subset \Omega$.

(In practice, extreme nonnormality may cause problems.)

Ex: Chebyshev iteration: Ω = interval or ellipse (sym. $\mathbb{R}$)

- If A is hpd CG implicitly constructs $\{p_n\} =: \{\pi_n\}$ s.t. π_ν = minimal polynomial (w.r.t. some Krylov space)
- If A is nonhermitian BCG (and CGS) do the same if no breakdown occurs.
  (But: nearbreakdown $\Rightarrow$ unstable)

Q: Is there another way to get information on the spectrum without computing it exactly?

- The recurrence coefficients of CG, which are related to recurrence coefficients $\alpha_n, \beta_n [, \gamma_n]$ of $\{\pi_n\}$ can be computed from the moments

$$\mu_k := \langle r_0, A^k r_0 \rangle := r_0^T A^k r_0$$

by the qd algorithm or the Chebyshev algorithm, but the mapping $\{\mu_k\} \longmapsto \{\alpha_k, \beta_k\}$ is ill-conditioned, and $A^k r_0$ is useless as approximation for z.

- If $\{r_n\}$ is sequence of residuals of some stable (but maybe not well converging) polynomial acc. method, one can use the modified moments

$$\nu_k := \langle r_0, r_n \rangle := r_0^T r_n$$

to compute $\{\alpha_k, \beta_k\}$ by the modified Chebyshev algorithm.

Does this work for nonhermitian A?