Preconditioners for Singular Black Box Matrices

William J. Turner

Department of Mathematics & Computer Science
Wabash College
Crawfordsville, IN 47933

ISSAC 2005 : Beijing, China
27 July 2005

WABASH COLLEGE

Outline

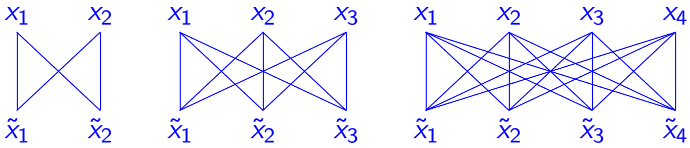
Arbitrary Radix Switches

Generalize butterfly switch to radix-ρ switch:

Arbitrary Radix Switches

Generalize butterfly switch to radix-ρ switch:

Example (Radix-ρ switches: $\rho = 2, 3, 4$)



Arbitrary Radix Switching Networks

Definition (ℓ-dimensional radix-ρ switching network)

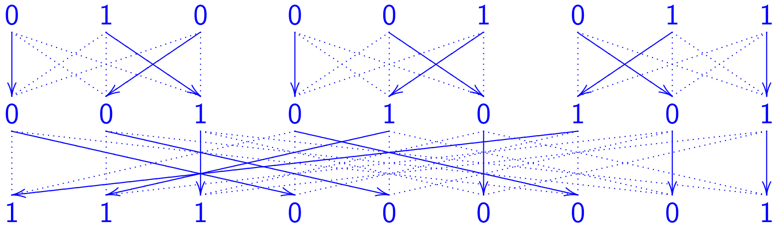
- Recursively defined
- $\rho^{\ell-1}$ radix-ρ switches to merge outputs of ρ radix-ρ subnetworks of dimension $\ell - 1$
- Merges ith output of each of the subnetworks

Arbitrary Radix Switching Networks

Definition (ℓ-dimensional radix-ρ switching network)

- Recursively defined
- $\rho^{\ell-1}$ radix-ρ switches to merge outputs of ρ radix-ρ subnetworks of dimension $\ell - 1$
- Merges ith output of each of the subnetworks

Example ($\rho = 3$, $\ell = 2$)



Arbitrary Radix Switching Networks

Lemma

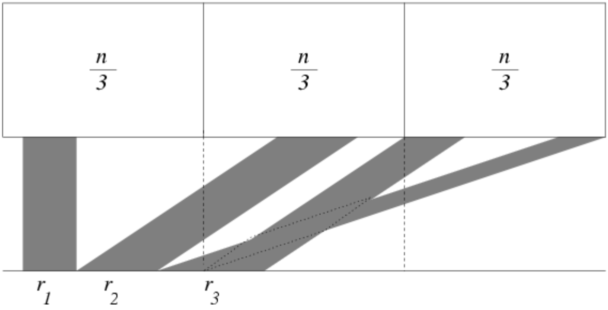
Let $n = \rho^\ell$ where $\rho \geq 2$. The ℓ-dimensional radix-ρ switching network can switch any r indices $1 \leq i_1 < \cdots < i_r \leq n$ into any desired contiguous block of indices. For our purposes, wrapping the block around the outside preserves contiguity.

Arbitrary Radix Switching Networks

Lemma

Let $n = \rho^\ell$ where $\rho \geq 2$. The ℓ-dimensional radix-ρ switching network can switch any r indices $1 \leq i_1 < \cdots < i_r \leq n$ into any desired contiguous block of indices. For our purposes, wrapping the block around the outside preserves contiguity.

Example ($\rho = 3$)



Generalized Arbitrary Radix Switching Networks

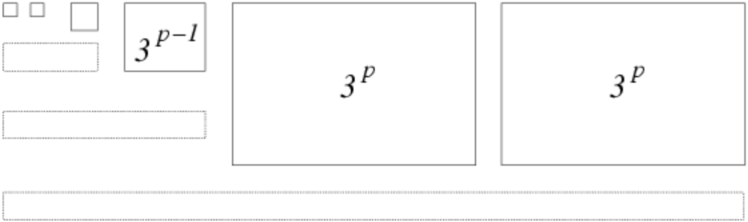
Definition (Generalized radix-ρ switching network)

- Let $n = \sum_{i=1}^{p} n_i$ where $\begin{cases} n_i = c_i \ \rho^{\ell_i} \\ c_i \in \{1, \ldots, \rho - 1\} \\ \ell_1 < \ell_2 < \cdots < \ell_p \end{cases}$
- Radix-$(c_k + 1)$ switches to merge $\sum_{i=1}^{k-1} n_i$ subnetwork with far right nodes of ρ^{ℓ_k} blocks
- Radix-c_k switches to merge other nodes of ρ^{ℓ_k} blocks

Generalized Arbitrary Radix Switching Networks

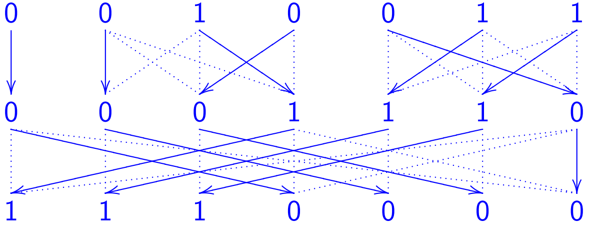
Definition (Generalized radix-ρ switching network)

- Let $n = \sum_{i=1}^{p} n_i$ where $\begin{cases} n_i = c_i \ \rho^{\ell_i} \\ c_i \in \{1, \ldots, \rho - 1\} \\ \ell_1 < \ell_2 < \cdots < \ell_p \end{cases}$
- Radix-$(c_k + 1)$ switches to merge $\sum_{i=1}^{k-1} n_i$ subnetwork with far right nodes of ρ^{ℓ_k} blocks
- Radix-c_k switches to merge other nodes of ρ^{ℓ_k} blocks



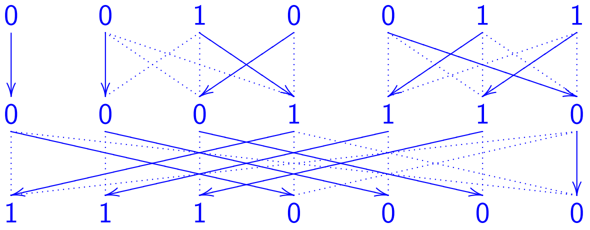
Generalized Arbitrary Radix Switching Networks

Example ($\rho = 3$, $n = 7 = 1\ \rho^0 + 2\ \rho^1$)



Generalized Arbitrary Radix Switching Networks

Example ($\rho = 3$, $n = 7 = 1\ \rho^0 + 2\ \rho^1$)



Theorem

Suppose $\rho \geq 2$. The generalized radix-ρ switching network described above can switch any r indices $1 \leq i_1 < \cdots < i_r \leq n$ into the contiguous block $1, 2, \ldots, r$. Furthermore, it has a depth of $\lceil \log_\rho(n) \rceil$ and a total of no more than $\rho^{\lceil \log_\rho(n) \rceil} \lceil \log_\rho(n) \rceil / \rho$ switches of radix at most ρ. The network attains this bound only when $n = \rho^{\lceil \log_\rho(n) \rceil}$ or $n < \rho$.

Building Preconditioners

- Replace each switch by $\rho \times \rho$ symbolic exchange matrix $\mathcal{S}_k$ to merge columns (not mix)
- Embed each exchange matrix as principal minor of $n \times n$ identity matrix to create $\hat{\mathcal{S}}_k$
- Symbolic preconditioner is product: $\mathcal{L} = \prod_{k=1}^{s} \hat{\mathcal{S}}_k$
- Randomly choose values for symbols to create probabilistic preconditioner: $\tilde{A} = AL$ where $L = \prod_{k=1}^{s} \hat{S}_k$

Building Preconditioners

Example ($\rho = 3$)

$$\mathcal{S}_k = \begin{bmatrix} \alpha_{1,1,k} & \alpha_{1,2,k} & \alpha_{1,3,k} \\ \alpha_{2,1,k} & \alpha_{2,2,k} & \alpha_{2,3,k} \\ \alpha_{3,1,k} & \alpha_{3,2,k} & \alpha_{3,3,k} \end{bmatrix}$$

$$\hat{\mathcal{S}}_k = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \alpha_{1,1,k} & & \alpha_{1,2,k} & & \alpha_{1,3,k} & & \\ & & & \ddots & & & & & \\ & & \alpha_{2,1,k} & & \alpha_{2,2,k} & & \alpha_{2,3,k} & & \\ & & & & & \ddots & & & \\ & & \alpha_{3,1,k} & & \alpha_{3,2,k} & & \alpha_{3,3,k} & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix}$$

Symbolic Toeplitz Matrix

$$\mathcal{T} = \begin{bmatrix} \alpha_\rho & \alpha_{\rho+1} & \dots & \alpha_{2\rho-1} \\ \alpha_{\rho-1} & \alpha_\rho & \dots & \alpha_{2\rho-2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1 & \alpha_2 & \dots & \alpha_\rho \end{bmatrix}$$

Symbolic Toeplitz Matrix

$$\mathcal{T} = \begin{bmatrix} \alpha_\rho & \alpha_{\rho+1} & \dots & \alpha_{2\rho-1} \\ \alpha_{\rho-1} & \alpha_\rho & \dots & \alpha_{2\rho-2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1 & \alpha_2 & \dots & \alpha_\rho \end{bmatrix}$$

Using the lexicographic monomial ordering with $\alpha_1 \prec \alpha_2 \prec \cdots \prec \alpha_{2n-1}$:

$$\mathrm{lt}\left(\det\left(\mathcal{T}_{[i_1,\dots,i_m;j_1,\dots,j_m]}\right)\right) = (-1)^{\lfloor m/2 \rfloor} \prod_{k=1}^{m} \alpha_{n+j_{m+1-k}-i_k}$$

Toeplitz Exchange Matrix

Let A' be submatrix of A affected by $\mathcal{T}$.

Toeplitz Exchange Matrix

Let A' be submatrix of A affected by $\mathcal{T}$.

- A' includes a subset of r' of the r linearly independent columns of A where $0 \leq r' \leq r$.

Toeplitz Exchange Matrix

Let A' be submatrix of A affected by $\mathcal{T}$.

- A' includes a subset of r' of the r linearly independent columns of A where $0 \leq r' \leq r$.
- $\det((A'\mathcal{T})_{[\mathscr{I},\mathscr{J}]}) = \sum_{\substack{\mathscr{K}=\{k_1,\ldots,k_m\} \\ 1 \leq k_1 < \cdots < k_m \leq n}} \det(A'_{[\mathscr{I},\mathscr{K}]}) \det(\mathcal{T}_{[\mathscr{K},\mathscr{J}]})$

Toeplitz Exchange Matrix

Let A' be submatrix of A affected by $\mathcal{T}$.

- A' includes a subset of r' of the r linearly independent columns of A where $0 \leq r' \leq r$.
- $\det((A'\mathcal{T})_{[\mathscr{I},\mathscr{J}]}) = \sum_{\substack{\mathscr{K}=\{k_1,\ldots,k_m\} \\ 1\leq k_1<\cdots<k_m\leq n}} \det(A'_{[\mathscr{I},\mathscr{K}]}) \det(\mathcal{T}_{[\mathscr{K},\mathscr{J}]})$
- Every set of $m \leq r'$ columns of $A'\mathcal{T}$ linearly independent

Toeplitz Exchange Matrix

Let A' be submatrix of A affected by $\mathcal{T}$.

- A' includes a subset of r' of the r linearly independent columns of A where $0 \leq r' \leq r$.
- $\det((A'\mathcal{T})_{[\mathscr{I},\mathscr{J}]}) = \sum_{\substack{\mathscr{K}=\{k_1,\ldots,k_m\} \\ 1\leq k_1<\cdots<k_m\leq n}} \det(A'_{[\mathscr{I},\mathscr{K}]}) \det(\mathcal{T}_{[\mathscr{K},\mathscr{J}]})$
- Every set of $m \leq r'$ columns of $A'\mathcal{T}$ linearly independent
- Embedded Toeplitz exchange matrix $\hat{\mathcal{T}}$ can move linear independence freely within one switch.

Toeplitz Exchange Matrix

Let A' be submatrix of A affected by $\mathcal{T}$.

- A' includes a subset of r' of the r linearly independent columns of A where $0 \leq r' \leq r$.
- $\det((A'\mathcal{T})_{[\mathscr{I},\mathscr{J}]}) = \sum_{\substack{\mathscr{K}=\{k_1,\ldots,k_m\} \\ 1 \leq k_1 < \cdots < k_m \leq n}} \det(A'_{[\mathscr{I},\mathscr{K}]}) \det(\mathcal{T}_{[\mathscr{K},\mathscr{J}]})$
- Every set of $m \leq r'$ columns of $A'\mathcal{T}$ linearly independent
- Embedded Toeplitz exchange matrix $\hat{\mathcal{T}}$ can move linear independence freely within one switch.
- Each $\hat{\mathcal{T}}_k$ uses different symbols $\Longrightarrow$ first r columns of $A(\prod \hat{\mathcal{T}}_k)$ linearly independent.

Toeplitz Exchange Matrix

Theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ have r linearly independent columns, let s be the number of switches in the generalized radix-ρ switching network, and let S be a finite subset of $\mathbb{F}$. Let N be the number of random numbers required in this network. If $a_1, \ldots, a_N$ are randomly chosen uniformly and independently from S, then the first r columns of $A(\prod_{k=1}^{s} \widehat{T}_k)$ are linearly independent with probability at least

$$1 - \frac{r\lceil \log_\rho(n) \rceil}{|S|} \geq 1 - \frac{n\lceil \log_\rho(n) \rceil}{|S|}.$$

Toeplitz Rank Preconditioner

- $\rho = n \Longrightarrow$ *any* r columns of $A\mathcal{T}$ linearly independent

Toeplitz Rank Preconditioner

- $\rho = n \Longrightarrow$ *any* r columns of $A\mathcal{T}$ linearly independent

- *Some* $r \times r$ principal minor of $A\mathcal{T}$ nonzero

Toeplitz Rank Preconditioner

- $\rho = n \Longrightarrow$ *any* r columns of $A\mathcal{T}$ linearly independent
- *Some* $r \times r$ principal minor of $A\mathcal{T}$ nonzero
- ATD preconditioner for Kaltofen-Saunders rank algorithm [Turner(2002), Ch.3]

Toeplitz Rank Preconditioner

- $\rho = n \Longrightarrow$ *any* r columns of $A\mathcal{T}$ linearly independent
- *Some* $r \times r$ principal minor of $A\mathcal{T}$ nonzero
- ATD preconditioner for Kaltofen-Saunders rank algorithm [Turner(2002), Ch.3]
- Do not need D!

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \sum_{m=0}^{n} (-1)^m E_m(A\mathcal{T})\lambda^{n-m}$
- $E_m(A\mathcal{T}) = \sum_{\substack{\mathscr{I}=\{i_1,\ldots,i_m\} \\ 1\leq i_1<\cdots<i_m\leq n}} \det((A\mathcal{T})_{[\mathscr{I},\mathscr{I}]})$

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \sum_{m=0}^{n} (-1)^m E_m(A\mathcal{T}) \lambda^{n-m}$
- $E_m(A\mathcal{T}) = \sum_{\substack{\mathscr{I}=\{i_1,\dots,i_m\} \\ 1 \le i_1 < \cdots < i_m \le n}} \det((A\mathcal{T})_{[\mathscr{I},\mathscr{I}]})$
- $m > r \Longrightarrow E_m(A\mathcal{T}) = 0$

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \sum_{m=0}^{n} (-1)^m E_m(A\mathcal{T}) \lambda^{n-m}$
- $E_m(A\mathcal{T}) = \sum_{\substack{\mathscr{I}=\{i_1,\dots,i_m\} \\ 1 \leq i_1 < \cdots < i_m \leq n}} \det((A\mathcal{T})_{[\mathscr{I},\mathscr{I}]})$
- $m > r \implies E_m(A\mathcal{T}) = 0$
- λ^{n-r} divides $\det(\lambda I - A\mathcal{T})$

Leading Term of Characteristic Polynomial

Let

- $\mathscr{P} = \{p_1, \ldots, p_r\}$ where $1 \leq p_1 < \cdots < p_r \leq n$ be the pivot column indices of A
- $\mathscr{Q} = \{q_1, \ldots, q_r\}$ where $1 \leq q_1 < \cdots < q_r \leq n$ be the indices of the last r linearly independent rows of A

Then, under the lexicographic monomial ordering with $\lambda \prec \alpha_1 \prec \alpha_2 \prec \cdots \prec \alpha_{2n-1}$,

$$\operatorname{lt}(\det(\lambda I - A\mathcal{T})) = (-1)^{\lfloor 3r/2 \rfloor} A_{[\mathscr{Q},\mathscr{P}]} \lambda^{n-r} \prod_{k=1}^{r} \alpha_{n+p_{m+1-k}-q_k}$$

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \lambda^{n-r} g$ where $\deg_\lambda(g) = r$ and λ does not divide g

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \lambda^{n-r} g$ where $\deg_\lambda(g) = r$ and λ does not divide g

- $\mathrm{lt}(g)$ constant in λ and at most linear in each α_i

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \lambda^{n-r} g$ where $\deg_\lambda(g) = r$ and λ does not divide g
- $\mathrm{lt}(g)$ constant in λ and at most linear in each α_i
- g squarefree

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \lambda^{n-r} g$ where $\deg_\lambda(g) = r$ and λ does not divide g
- $\mathrm{lt}(g)$ constant in λ and at most linear in each α_i
- g squarefree
- $f^{A\mathcal{T}} = \lambda^p g$ where $p \geq 1$

Characteristic Polynomial

- $\det(\lambda I - A\mathcal{T}) = \lambda^{n-r} g$ where $\deg_\lambda(g) = r$ and λ does not divide g
- $\mathrm{lt}(g)$ constant in λ and at most linear in each α_i
- g squarefree
- $f^{A\mathcal{T}} = \lambda^p g$ where $p \geq 1$
- $\deg_\lambda(f^{A\mathcal{T}}) \geq r + 1$

Degree Upper Bound

- Sylvester's inequality: $\text{rank}(A\mathcal{T}) \leq r$

Degree Upper Bound

- Sylvester's inequality: $\text{rank}(A\mathcal{T}) \leq r$
- λ divides the $n - r$ largest invariant factors $s_k(\lambda I - A\mathcal{T})$

Degree Upper Bound

- Sylvester's inequality: $\text{rank}(A\mathcal{T}) \leq r$
- λ divides the $n - r$ largest invariant factors $s_k(\lambda I - A\mathcal{T})$
- $\deg(f^{A\mathcal{T}}) \leq r + 1$

Degree Upper Bound

- Sylvester's inequality: $\mathrm{rank}(A\mathcal{T}) \leq r$
- λ divides the $n - r$ largest invariant factors $s_k(\lambda I - A\mathcal{T})$
- $\deg(f^{A\mathcal{T}}) \leq r + 1$
- $\deg(f^{A\mathcal{T}}) = r + 1$

Toeplitz Rank Preconditioner

Theorem

Let $\mathbb{F}$ *be a field,* $A \in \mathbb{F}^{n\times n}$ *have rank* r *with* $r \le n-1$*, and* S *be a finite subset of* $\mathbb{F}$*. If*

$$T = \begin{bmatrix} a_n & a_{n+1} & \dots & a_{2n-1} \\ a_{n-1} & a_n & \dots & a_{2n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_n \end{bmatrix} \in \mathbb{F}^{n\times n},$$

where $a_1, \dots, a_{2n-1}$ *are chosen uniformly and independently from* S*, then the matrix* AT *has characteristic polynomial* $\det(\lambda I - AT) = \lambda^{n-r} g(\lambda)$ *and minimal polynomial* $f^{AT} = \lambda\, g(\lambda)$ *where* $g(0) \neq 0$ *and* $\deg(f^{AT}) = r+1$*, all with probability at least*

$$1 - \frac{r(r+1)}{2|S|} \ge 1 - \frac{n(n-1)}{2|S|}.$$

Conclusion

- Generalization of Chen et al. butterfly preconditioners to arbitrary radix switches

Conclusion

- Generalization of Chen et al. butterfly preconditioners to arbitrary radix switches
  - Small increase in probability bound for success
  - $\rho \longrightarrow n$ allows search for other rank preconditioners

Conclusion

- Generalization of Chen et al. butterfly preconditioners to arbitrary radix switches
  - Small increase in probability bound for success
  - $\rho \longrightarrow n$ allows search for other rank preconditioners
- Monomial ordering and leading term arguments

Conclusion

- Generalization of Chen et al. butterfly preconditioners to arbitrary radix switches
  - Small increase in probability bound for success
  - $\rho \longrightarrow n$ allows search for other rank preconditioners
- Monomial ordering and leading term arguments
  - Toeplitz matrices as arbitrary radix exchange matrices
  - Toeplitz matrix as rank preconditioner
    - Better probability of success
    - Asymptotically slower than butterfly preconditioner
  - Apply to other matrices?

References

L. Chen, W. Eberly, E. Kaltofen, B. D. Saunders, W. J. Turner, & G. Villard (2002).
Efficient Matrix Preconditioners for Black Box Linear Algebra.
Linear Algebra and its Applications, 343-344:119–146.
Special issue on *Infinite Systems of Linear Equations Finitely Specified*.

W. J. Turner (2002).
Black Box Linear Algebra with the LinBox Library.
Ph.D. thesis, Raleigh, North Carolina.