

Accurate Solutions of Structured Linear Systems

Juan M. Molera and Froilán M. Dopico

Universidad Carlos III de Madrid

SIAM LA09, Monterey Bay-Seaside, October 26-29, 2009

Problem

Given a **structured (Cauchy, Vandermonde, ...)** invertible matrix $A \in \mathbb{R}^{n \times n}$, is it possible to solve a linear system

$$Az = b$$

with **high relative accuracy (hra)**?

Problem

Given a **structured (Cauchy, Vandermonde, ...)** invertible matrix $A \in \mathbb{R}^{n \times n}$, is it possible to solve a linear system

$$Az = b$$

with **high relative accuracy (hra)**?

- Until now, **only possible in the Totally Positive case**
 - Vandermonde [1970, Björk&Pereyra; 1987, Higham]
 - Cauchy [1989, Gohberg&Koltracht; 1999, Boros&Kailath&Olshevsky]

Problem

Given a **structured (Cauchy, Vandermonde, ...)** invertible matrix $A \in \mathbb{R}^{n \times n}$, is it possible to solve a linear system

$$Az = b$$

with **high relative accuracy (hra)**?

- Until now, **only possible in the Totally Positive case**
 - Vandermonde [1970, Björk&Pereyra; 1987, Higham]
 - Cauchy [1989, Gohberg&Koltracht; 1999, Boros&Kailath&Olshevsky]
- Here we will present a method that **gets (hra) for any Cauchy or Vandermonde (and many other classes)**.

Problem

Given a **structured (Cauchy, Vandermonde, ...)** invertible matrix $A \in \mathbb{R}^{n \times n}$, is it possible to solve a linear system

$$Az = b$$

with **high relative accuracy (hra)**?

- Until now, **only possible in the Totally Positive case**
 - Vandermonde [1970, Björk&Pereyra; 1987, Higham]
 - Cauchy [1989, Gohberg&Koltracht; 1999, Boros&Kailath&Olshevsky]
- Here we will present a method that **gets (hra) for any Cauchy or Vandermonde (and many other classes)**. It is based on **accurate Rank Revealing Decompositions**, that are known to exist for many classes of structured matrices.

High relative accuracy

We say that the solution of

$$Az = b$$

is computed with **high relative accuracy (hra)** if

$$\frac{\|\hat{z} - z\|}{\|z\|} = O(\mathbf{u})$$

with \hat{z} the computed solution and \mathbf{u} the unit roundoff,
independently of the condition number of A .

Rank Revealing Decompositions

- Rank Revealing Decompositions have been widely used to compute
 - Singular Values and Vectors [1999, Demmel et al.]
 - Eigenvalues and Eigenvectors [2003, Dopico&M&Moro; 2005, Dopico&Koev; 2009 Dopico&Koev&M]

Rank Revealing Decompositions

- Rank Revealing Decompositions have been widely used to compute
 - Singular Values and Vectors [1999, Demmel et al.]
 - Eigenvalues and Eigenvectors [2003, Dopico&M&Moro; 2005, Dopico&Koev; 2009 Dopico&Koev&M]
- In this work **RRDs will be used to solve linear systems.**

Rank Revealing Decompositions

- Rank Revealing Decompositions have been widely used to compute
 - Singular Values and Vectors [1999, Demmel et al.]
 - Eigenvalues and Eigenvectors [2003, Dopico&M&Moro; 2005, Dopico&Koev; 2009 Dopico&Koev&M]
- In this work **RRDs will be used to solve linear systems.**
- The algorithm presented has $O(n^3)$ complexity.

Cauchy and Vandermonde matrices

Given $x, y \in \mathbb{R}^n$, a **Cauchy matrix** is given by

$$C_{i,j} = \frac{1}{x_i + y_j}$$

$$C(x, y) = \begin{bmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \cdots & \frac{1}{x_1+y_n} \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \cdots & \frac{1}{x_2+y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \cdots & \frac{1}{x_n+y_n} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

$C(x, y)$ is **Totally Positive** if

$$x_1 + y_1 > 0$$

$$x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n$$

$$y_1 \leq y_2 \leq \dots \leq y_{n-1} \leq y_n$$

Cauchy and Vandermonde matrices

Given $x \in \mathbb{R}^n$, a **Vandermonde matrix** is given by

$$V_{i,j} = x_j^{i-1}$$

$$V(x) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

$V(x)$ is **Totally Positive** if

$$0 < x_1 < x_2 < \cdots < x_n$$

Cauchy and Vandermonde matrices: Properties

- They appear in many applications.
- They have many notable algebraic properties.

Cauchy and Vandermonde matrices: Properties

- They appear in many applications.
- They have many notable algebraic properties.
- They have huge condition numbers.

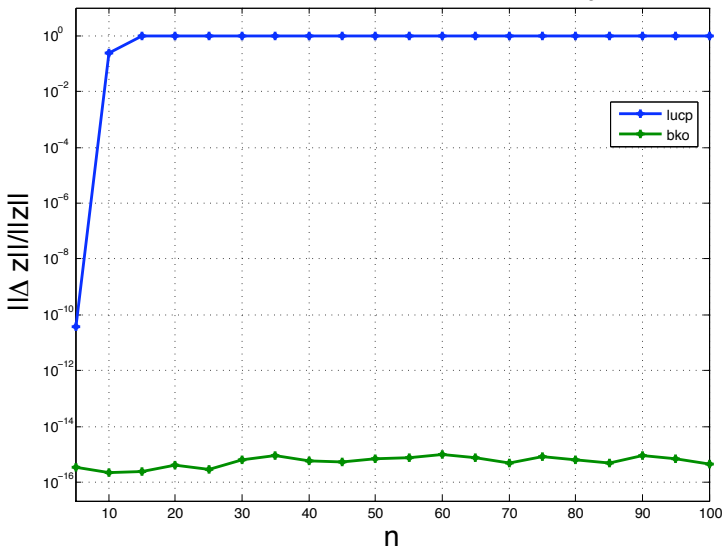
Cauchy and Vandermonde matrices: Properties

- They appear in many applications.
- They have many notable algebraic properties.
- They have huge condition numbers.
- They are related through a DFT.
- There exist fast $O(n^2)$ algorithms to solve linear systems. [1970, Björk&Pereyra; 1999, Boros&Kailath&Olshevsky]
- For these fast algorithms [hra](#) is achieved (guaranteed for certain rhs) when they are totally positive.

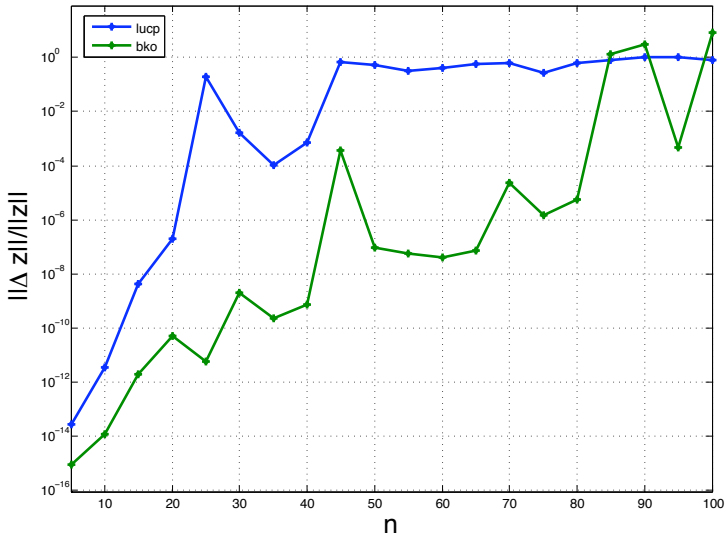
Cauchy and Vandermonde matrices: Properties

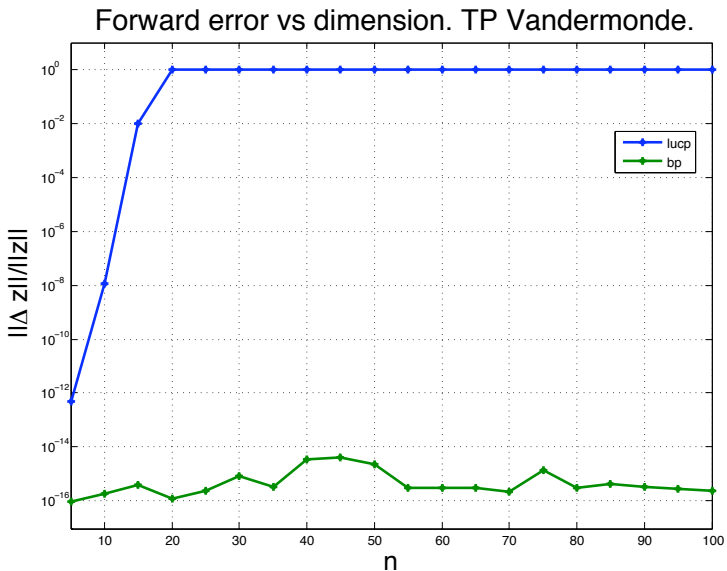
- They appear in many applications.
- They have many notable algebraic properties.
- They have huge condition numbers.
- They are related through a DFT.
- There exist fast $O(n^2)$ algorithms to solve linear systems. [1970, Björk&Pereyra; 1999, Boros&Kailath&Olshevsky]
- For these fast algorithms hra is achieved (guaranteed for certain rhs) when they are totally positive.
- Backward errors related to the perturbation of the rhs. [1987, Higham; 1999, Boros&Kailath&Olshevsky]

Forward error vs dimension. TP Cauchy matrices.

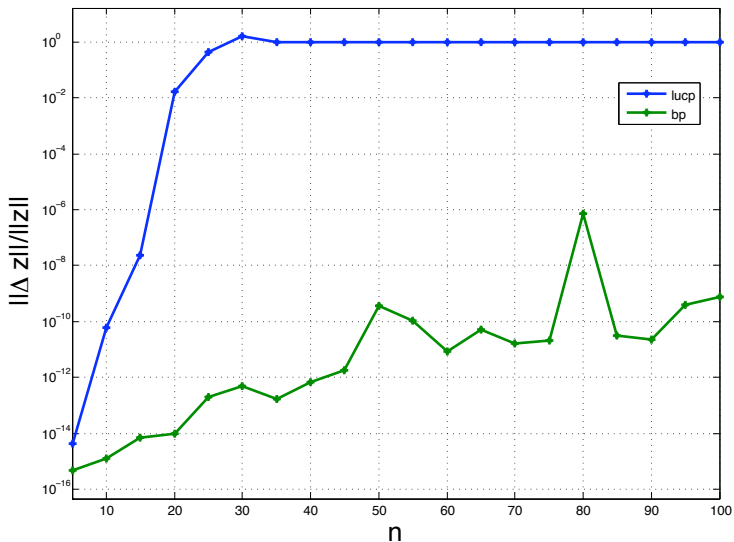


Forward error vs dimension. Random Cauchy matrices.





Forward error vs dimension. Random Vandermonde.



Perturbing the right-hand side

If $A \in \mathbb{R}^{n \times n}$ is invertible and

$$Az = b \quad \text{and} \quad A(z + \Delta z) = b + \Delta b$$

then

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|A^{-1}\|_2 \|\Delta b\|_2}{\|z\|_2} = \kappa_2(A, b) \frac{\|\Delta b\|_2}{\|b\|_2}$$

with

$$\kappa_2(A, b) \equiv \frac{\|A^{-1}\|_2 \|b\|_2}{\|z\|_2} = \frac{\|b\|_2}{\sigma_n \|z\|_2}$$

being σ_n the smallest singular value of A .

Effective conditioning

Theorem (Chan & Foulser, 1988)

If

$$A = U\Sigma V^T = [\tilde{U} \ U_k] \begin{bmatrix} \tilde{\Sigma} & 0 \\ 0 & \Sigma_k \end{bmatrix} \begin{bmatrix} \tilde{V} \\ V_k \end{bmatrix}$$

is a (partitioned) SVD of A , and $Az = b$, then

$$\kappa_2(A, b) = \frac{\|b\|_2}{\sigma_n \|z\|_2} \leq \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2}, \quad \forall k = 1 : n$$

where $P_k = U_k U_k^T$ is the orthogonal projector onto the subspace spanned by the last k left singular vectors of A .

Effective conditioning

For the perturbation problem $A(z + \Delta z) = b + \Delta b$ this means

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \kappa_2(A, b) \frac{\|\Delta b\|_2}{\|b\|_2} \leq \kappa_{eff}(A, b) \frac{\|\Delta b\|_2}{\|b\|_2}$$

with

$$\kappa_{eff}(A, b) = \min_{k=1:n} \left\{ \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2} \right\} \leq \frac{\|b\|_2}{\|P_1 b\|_2}$$

Effective conditioning

For the perturbation problem $A(z + \Delta z) = b + \Delta b$ this means

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \kappa_2(A, b) \frac{\|\Delta b\|_2}{\|b\|_2} \leq \kappa_{eff}(A, b) \frac{\|\Delta b\|_2}{\|b\|_2}$$

with

$$\kappa_{eff}(A, b) = \min_{k=1:n} \left\{ \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2} \right\} \leq \frac{\|b\|_2}{\|P_1 b\|_2}$$

Notice that, except in an unlikely case where $u_n^T b \approx 0$, we will have

$$\|P_1 b\|_2 = u_n^T b \lesssim \|b\|_2$$

with u_n the last left singular vector of A .

Effective conditioning

For the perturbation problem $A(z + \Delta z) = b + \Delta b$ this means

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \kappa_2(A, b) \frac{\|\Delta b\|_2}{\|b\|_2} \leq \kappa_{eff}(A, b) \frac{\|\Delta b\|_2}{\|b\|_2}$$

with

$$\kappa_{eff}(A, b) = \min_{k=1:n} \left\{ \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2} \right\} \leq \frac{\|b\|_2}{\|P_1 b\|_2}$$

If, for example, $\|P_1 b\|_2 = u_n^T b \geq \frac{\|b\|_2}{n}$, then

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq n \frac{\|\Delta b\|_2}{\|b\|_2}$$

Algorithm

- **Input:** $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$
- **Output:** z solution of $Az = b$
- **Step 1:** Compute an accurate RRD of $A = XDY^T$
- **Step 2:** Solve the three systems

$$Xs = b \quad \longrightarrow \quad s$$

$$Dw = s \quad \longrightarrow \quad w$$

$$Yz = w \quad \longrightarrow \quad z$$

Rank Revealing Decomposition

- 1 Given $A \in \mathbb{C}^{m \times n}$, with $\text{rank}(A) = r$,

$$A = XDY^T,$$

with $X \in \mathbb{C}^{m \times r}$, $Y \in \mathbb{C}^{n \times r}$ and $D \in \mathbb{C}^{r \times r}$ diagonal, it is called a **Rank Revealing Decomposition (RRD)** of A if

$$\kappa(X), \kappa(Y) \gtrsim 1 \quad \text{and} \quad \kappa(D) \approx \kappa(A).$$

Rank Revealing Decomposition

- 1 Given $A \in \mathbb{C}^{m \times n}$, with $\text{rank}(A) = r$,

$$A = XDY^T,$$

with $X \in \mathbb{C}^{m \times r}$, $Y \in \mathbb{C}^{n \times r}$ and $D \in \mathbb{C}^{r \times r}$ diagonal, it is called a **Rank Revealing Decomposition (RRD)** of A if

$$\kappa(X), \kappa(Y) \gtrsim 1 \quad \text{and} \quad \kappa(D) \approx \kappa(A).$$

- 2 We say that an RRD, $A = XDY^T$, is **accurate** if the computed factors \hat{X} , \hat{Y} and \hat{D} obey

$$\frac{\|\hat{X} - X\|}{\|X\|} = O(\mathbf{u}), \quad \frac{\|\hat{Y} - Y\|}{\|Y\|} = O(\mathbf{u}) \quad \text{and} \quad \frac{|\hat{D} - D|}{|D|} = O(\mathbf{u})$$

Structured matrices and RRDs

- **RRDs** can be computed in practice from Gaussian Elimination with Complete Pivoting (GECP) (i.e., direct methods) .

Structured matrices and RRDs

- **RRDs** can be computed in practice from Gaussian Elimination with Complete Pivoting (GECP) (i.e., direct methods).
- But **accurate RRDs** can only be computed for matrices with special structures:
 - Scaled-Cauchy, Vandermonde (DFT + GECP). [Demmel]
 - Diagonally Dominant M-Matrices. [Demmel and Koev]
 - Polynomial Vandermonde. [Demmel and Koev]
 - Well Scaled Positive Definite. [Demmel and Veselić]
 - Acyclic Matrices (include bidiagonal). [Demmel and Gragg]
 - Diagonally Dominant. [Qiang Ye, Dopico and Koev]
 - DSTU. [Demmel]
 - Graded Matrices. [Dopico, M]
 - ...

Backward Error

Theorem

If \hat{z} is the computed solution of $XDY^T z = b$ using previous Algorithm, with roundoff error \mathbf{u} , then

$$(X + \Delta X)(D + \Delta D)(Y + \Delta Y)^T \hat{z} = b$$

with

$$\|\Delta X\| \leq O(\mathbf{u})\|X\|, \quad \|\Delta Y\| \leq O(\mathbf{u})\|Y\|, \quad |\Delta D| \leq O(\mathbf{u})|D|$$

RRD Perturbation Theory

Theorem

Let $A = XDY^T \in \mathbb{R}^{n \times n}$ be an RRD of A and z the solution of $XDY^T z = b$, then the solution of

$$(X + \Delta X)(D + \Delta D)(Y + \Delta Y)^T(z + \Delta z) = b$$

obeys, if $\|\Delta X\| \leq \epsilon \|X\|$, $\|\Delta Y\| \leq \epsilon \|Y\|$, $|\Delta D| \leq \epsilon |D|$,

RRD Perturbation Theory

Theorem

Let $A = XDY^T \in \mathbb{R}^{n \times n}$ be an RRD of A and z the solution of $XDY^T z = b$, then the solution of

$$(X + \Delta X)(D + \Delta D)(Y + \Delta Y)^T(z + \Delta z) = b$$

obeys, if $\|\Delta X\| \leq \epsilon \|X\|$, $\|\Delta Y\| \leq \epsilon \|Y\|$, $|\Delta D| \leq \epsilon |D|$,

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|b\|_2}{\sigma_n \|z\|_2} [2\kappa_2(X) + \kappa_2(Y)] O(\epsilon)$$

with $\kappa_2(X) = \|X^{-1}\|_2 \|X\|_2$, $\kappa_2(Y) = \|Y^{-1}\|_2 \|Y\|_2$.

RRD Perturbation Theory

Theorem

Let $A = XDY^T \in \mathbb{R}^{n \times n}$ be an RRD of A and z the solution of $XDY^T z = b$, then the solution of

$$(X + \Delta X)(D + \Delta D)(Y + \Delta Y)^T(z + \Delta z) = b$$

obeys, if $\|\Delta X\| \leq \epsilon \|X\|$, $\|\Delta Y\| \leq \epsilon \|Y\|$, $|\Delta D| \leq \epsilon |D|$,

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|b\|_2}{\sigma_n \|z\|_2} [2\kappa_2(X) + \kappa_2(Y)] O(\epsilon)$$

with $\kappa_2(X) = \|X^{-1}\|_2 \|X\|_2$, $\kappa_2(Y) = \|Y^{-1}\|_2 \|Y\|_2$.

Forward Error

Theorem

If \hat{z} is the computed solution of $Az = b$ using previous Algorithm, with roundoff error \mathbf{u} , then

Forward Error

Theorem

If \hat{z} is the computed solution of $Az = b$ using previous Algorithm, with roundoff error \mathbf{u} , then

$$\frac{\|\hat{z} - z\|_2}{\|z\|_2} \leq \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2} \left[2\kappa_2(\hat{X}) + \kappa_2(\hat{Y}) \right] O(\mathbf{u})$$

Forward Error

Theorem

If \hat{z} is the computed solution of $Az = b$ using previous Algorithm, with roundoff error \mathbf{u} , then

$$\frac{\|\hat{z} - z\|_2}{\|z\|_2} \leq \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2} \left[2\kappa_2(\hat{X}) + \kappa_2(\hat{Y}) \right] O(\mathbf{u})$$

and, in particular, if $|u_n^T b| \geq \|b\|_2/n$,

Forward Error

Theorem

If \hat{z} is the computed solution of $Az = b$ using previous Algorithm, with roundoff error \mathbf{u} , then

$$\frac{\|\hat{z} - z\|_2}{\|z\|_2} \leq \frac{\sigma_{n+1-k}}{\sigma_n} \frac{\|b\|_2}{\|P_k b\|_2} \left[2\kappa_2(\hat{X}) + \kappa_2(\hat{Y}) \right] O(\mathbf{u})$$

and, in particular, if $|u_n^T b| \geq \|b\|_2/n$,

$$\frac{\|\hat{z} - z\|_2}{\|z\|_2} \leq n \left[2\kappa_2(\hat{X}) + \kappa_2(\hat{Y}) \right] O(\mathbf{u})$$

being u_n the last left singular vector of A .

Numerical Experiments: Cauchy matrices

We have built Cauchy matrices with

- TP matrices (ordered x and y vectors) and non TP matrices (random x and y vectors)
- Random rhs vector b
- Sizes 5:5:100
- $10^6 \lesssim \kappa(C(x, y)) \lesssim 10^{200}$

Numerical Experiments: Cauchy matrices

We have built Cauchy matrices with

- TP matrices (ordered x and y vectors) and non TP matrices (random x and y vectors)
- Random rhs vector b
- Sizes 5:5:100
- $10^6 \lesssim \kappa(C(x, y)) \lesssim 10^{200}$

We have used four algorithms

- GECP
- BKO fast ($O(n^2)$) algorithm (Boros-Kailath-Olshevsky)
- GS-Cauchy [BKO] fast algorithm with pre-pivoting (Rational Leja ordering)
- Rank Revealing ($O(n^3)$) algorithm (Demmel)+ MATLAB ($Y \setminus D \setminus X \setminus b$)

Computable bound

We have checked the bound in the forward error

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|b\|_2}{\sigma_n \|z\|_2} [2\kappa_2(X) + \kappa_2(Y)] O(\mathbf{u})$$

Computable bound

We have checked the bound in the forward error

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|b\|_2}{\sigma_n \|z\|_2} [2\kappa_2(X) + \kappa_2(Y)] O(\mathbf{u})$$

Computable bound

We have checked the bound in the forward error

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|b\|_2}{\sigma_n \|z\|_2} [2\kappa_2(X) + \kappa_2(Y)] O(\mathbf{u})$$

Using

$$\frac{1}{\sigma_n} = \|A^{-1}\|_2 = \|Y^{-T} D^{-1} X^{-1}\|_2 \leq \frac{\|Y^{-T} X^{-1}\|_2}{|d_n|}$$

Computable bound

We have checked the bound in the forward error

$$\frac{\|\Delta z\|_2}{\|z\|_2} \leq \frac{\|b\|_2}{\sigma_n \|z\|_2} [2\kappa_2(X) + \kappa_2(Y)] O(\mathbf{u})$$

Using

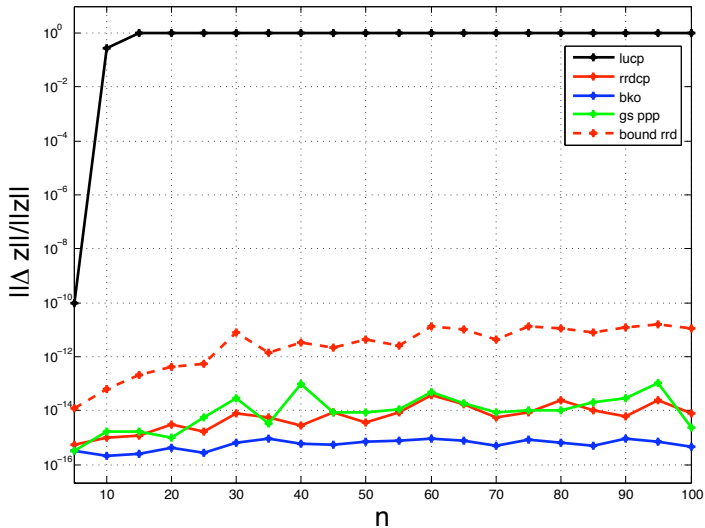
$$\frac{1}{\sigma_n} = \|A^{-1}\|_2 = \|Y^{-T} D^{-1} X^{-1}\|_2 \leq \frac{\|Y^{-T} X^{-1}\|_2}{|d_n|}$$

we get **the computable approximation**

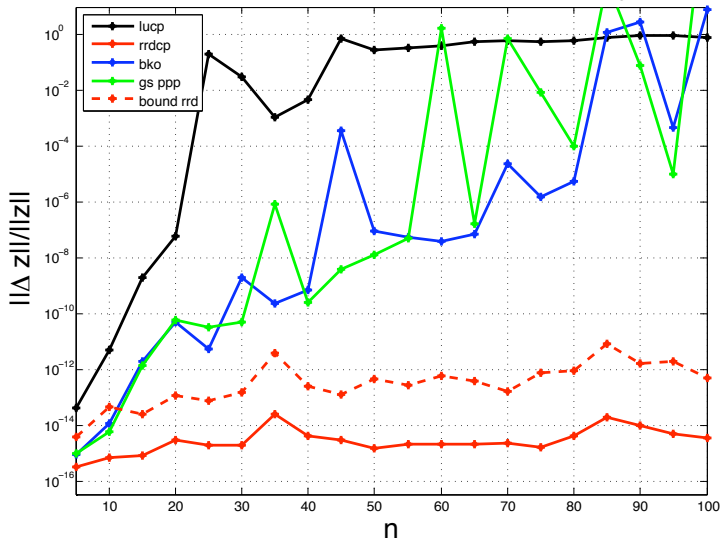
$$\text{bound_rrd} = \frac{\|\hat{Y}^{-1}\|_2 \|\hat{X}^{-1}\|_2 \|b\|_2}{|d_n| \|\hat{z}\|_2} [2\kappa_2(\hat{X}) + \kappa_2(\hat{Y})] \mathbf{u}$$

independent of σ_n or A^{-1} .

Forward error vs dimension. TP Cauchy matrices.



Forward error vs dimension. Random Cauchy matrices.



Numerical Experiments: Vandermonde matrices

We have built Vandermonde matrices

- TP matrices (ordered x vector) and non TP matrices (random x vector)
- Random rhs vector b
- Sizes 5:5:100
- $10^2 \lesssim \kappa(C(x, y)) \lesssim 10^{100}$

Numerical Experiments: Vandermonde matrices

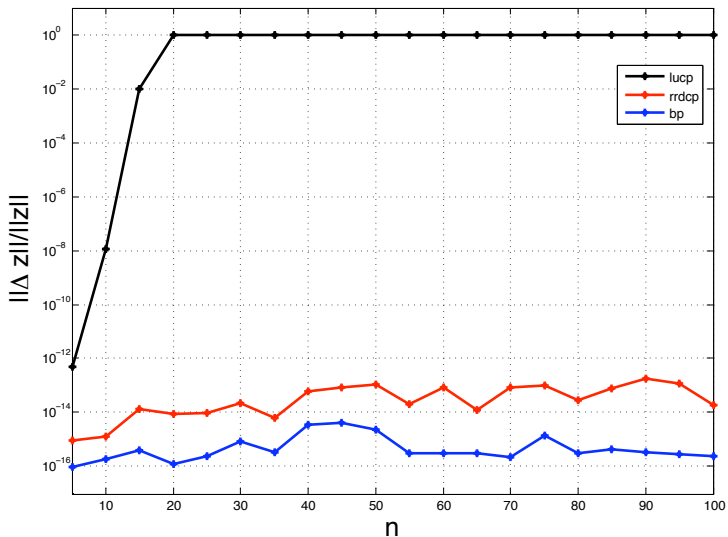
We have built Vandermonde matrices

- **TP matrices** (ordered x vector) and **non TP matrices** (random x vector)
- Random rhs vector b
- Sizes 5:5:100
- $10^2 \lesssim \kappa(C(x, y)) \lesssim 10^{100}$

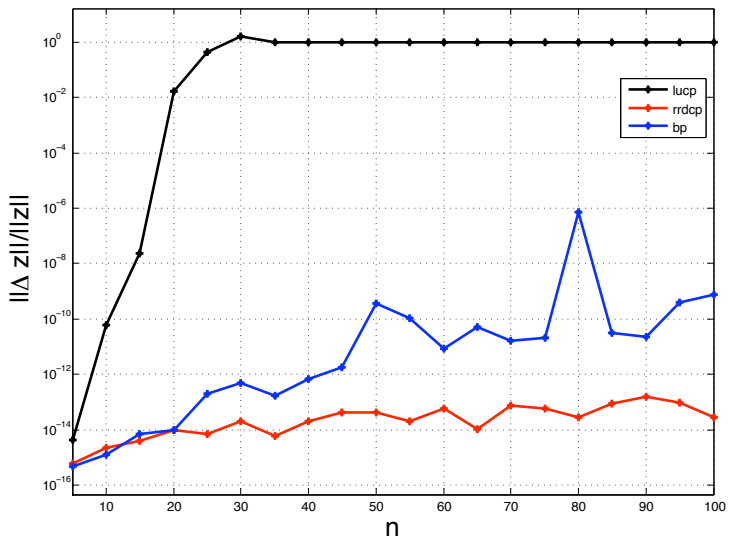
We have used three algorithms

- GECP
- Björk-Pereyra fast ($O(n^2)$) algorithm.
- Rank Revealing ($O(n^3)$) algorithm (Demmel)+ MATLAB ($Y \setminus D \setminus X \setminus b$)

Forward error vs dimension. TP Vandermonde.



Forward error vs dimension. Random Vandermonde.



Conclusions

- We have computed with relative accuracy the solution of structured linear systems, even if they are very badly conditioned.
- We have used an $O(n^3)$ algorithm that uses an accurate RRD.
- We have shown satisfactory numerical results for Cauchy and Vandermonde matrices.
- It will work for any kind of matrices for which accurate RRDs can be computed, now or in the future, and for, almost, any b .

Conclusions

- We have computed with relative accuracy the solution of structured linear systems, even if they are very badly conditioned.
- We have used an $O(n^3)$ algorithm that uses an accurate RRD.
- We have shown satisfactory numerical results for Cauchy and Vandermonde matrices.
- It will work for any kind of matrices for which accurate RRDs can be computed, now or in the future, and for, almost, any b .

Future work

Is it possible to do it with an $O(n^2)$, BP or BKO type, algorithm?

Conclusions

- We have computed with relative accuracy the solution of structured linear systems, even if they are very badly conditioned.
- We have used an $O(n^3)$ algorithm that uses an accurate RRD.
- We have shown satisfactory numerical results for Cauchy and Vandermonde matrices.
- It will work for any kind of matrices for which accurate RRDs can be computed, now or in the future, and for, almost, any b .

Future work

Is it possible to do it with an $O(n^2)$, BP or BKO type, algorithm?

Thank you for your attention.

