# Conditioning of the Generalized Hankel Eigenvalue Problem

George Labahn

Symbolic Computation Group
David R Cheriton School of Computer Science
University of Waterloo, Canada

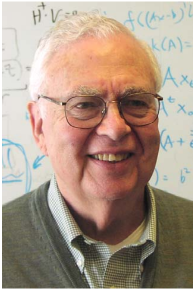
Golub Day : Feb 29, 2008

# Outline

# Co-authors

**Bernd Beckermann**

**Gene Golub**
**(1932-2007)**



**B. Beckermann, G. Golub, G. Labahn,**
**On the numerical condition of a generalized Hankel eigenvalue problem, *Numerische Mathematik*, 106 (2007) 41-68**

# The General Problem

**Given :** $h_0, h_1, ..., h_{2n-1}$ (moments),

**Find :** $c_1, \ldots, c_n$ and $\lambda_1, \ldots, \lambda_n$ such that

$$\begin{array}{lcccccccc} h_0 & = & c_1 & + & c_2 & + & \cdots & + & c_n \\ h_1 & = & c_1\lambda_1 & + & c_2\lambda_2 & + & \cdots & + & c_n\lambda_n \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ h_{2n-1} & = & c_1{\lambda_1}^{2n-1} & + & c_2{\lambda_2}^{2n-1} & + & \cdots & + & c_n{\lambda_n}^{2n-1} \end{array}$$

Correct numbers of unknowns but problem nonlinear.

# The Specific Problem + Results

Assume the $h_0, h_1, \ldots, h_{2n-1}$ are all approximately known.

Our question : What about the sensitivity of

$$G_n \quad : \quad \mathbb{C}^{2n} \mapsto \mathbb{C}^n$$
$$(h_0, ..., h_{2n-1}) \mapsto (\lambda_1, ..., \lambda_n)?$$

- Result 1 : Structured error = Unstructured error
- Result 2 : In general conditioning is terrible

  But in some cases randomization can help (Giesbrecht-Labahn-Lee).

# The Specific Problem + Results

Assume the $h_0, h_1, \ldots, h_{2n-1}$ are all approximately known.

Our question : What about the sensitivity of

$$\begin{aligned} G_n \quad : \quad & \mathbb{C}^{2n} \mapsto \mathbb{C}^n \\ & (h_0, ..., h_{2n-1}) \mapsto (\lambda_1, ..., \lambda_n)? \end{aligned}$$

- Result 1 : Structured error = Unstructured error
- Result 2 : In general conditioning is terrible

  But in some cases randomization can help (Giesbrecht-Labahn-Lee).

# The Specific Problem + Results

Assume the $h_0, h_1, \ldots, h_{2n-1}$ are all approximately known.

Our question : What about the sensitivity of

$$\begin{aligned} G_n \quad : \quad & \mathbb{C}^{2n} \mapsto \mathbb{C}^n \\ & (h_0, ..., h_{2n-1}) \mapsto (\lambda_1, ..., \lambda_n)? \end{aligned}$$

- Result 1 : Structured error = Unstructured error
- Result 2 : In general conditioning is terrible

  But in some cases randomization can help (Giesbrecht-Labahn-Lee).

# Tools used:

For error analysis :

- Formulae for inverses of Vandermonde matrices (Lagrange polynomials)
- Singular values of tri-diagonal matrices

For sensitivity analysis :

- Condition numbers for row scaled Vandermonde matrices
- Techniques from optimization theory and potential theory

# Motivation : Sparse Interpolation

Given a black box representation of a multivariate polynomial $p$ :

$$\vec{x} \longrightarrow \boxed{\phantom{xxxx}} \longrightarrow p(\vec{x})$$

Find powers $\vec{m}_1, \ldots, \vec{m}_n$ and coefficients $c_1, \ldots, c_n$ so that

$$p(x_1, \ldots, x_d) = c_1 \mathbf{x}^{\vec{m}_1} + \cdots + c_n \mathbf{x}^{\vec{m}_n}$$

where $\mathbf{x}^{\vec{m}_i} = x_1^{m_{1,i}} \cdots x_d^{m_{d,i}}$.

- In symbolic case : Ben-Or/Tiwari (1988).
- Basically uses a variation of Prony's method (1795).
- In numeric environment : Giesbrecht-Labahn-Lee (2006).

# Motivation : Sparse Interpolation

Given a black box representation of a multivariate polynomial $p$ :


Find powers $\vec{m}_1, \ldots, \vec{m}_n$ and coefficients $c_1, \ldots, c_n$ so that

$$p(x_1, \ldots, x_d) \quad = \quad c_1 \mathbf{x}^{\vec{m}_1} + \cdots + c_n \mathbf{x}^{\vec{m}_n}$$

where $\mathbf{x}^{\vec{m}_i} = x_1^{m_{1,i}} \cdots x_d^{m_{d,i}}$.

- In symbolic case : Ben-Or/Tiwari (1988).
- Basically uses a variation of Prony's method (1795).
- In numeric environment : Giesbrecht-Labahn-Lee (2006).

# Motivation : Sparse Interpolation

Given a black box representation of a multivariate polynomial $p$ :

$\vec{x} \longrightarrow$  $\longrightarrow p(\vec{x})$

Find powers $\vec{m}_1, \ldots, \vec{m}_n$ and coefficients $c_1, \ldots, c_n$ so that

$$p(x_1, \ldots, x_d) = c_1 \mathbf{x}^{\vec{m}_1} + \cdots + c_n \mathbf{x}^{\vec{m}_n}$$

where $\mathbf{x}^{\vec{m}_i} = x_1^{m_{1,i}} \cdots x_d^{m_{d,i}}$.

- In symbolic case : Ben-Or/Tiwari (1988).
- Basically uses a variation of Prony's method (1795).
- In numeric environment : Giesbrecht-Labahn-Lee (2006).

# Motivation : Sparse Interpolation

Given a black box representation of a multivariate polynomial $p$ :

$\vec{x} \longrightarrow$  $\longrightarrow p(\vec{x})$

Find powers $\vec{m}_1, \ldots, \vec{m}_n$ and coefficients $c_1, \ldots, c_n$ so that

$$p(x_1, \ldots, x_d) \quad = \quad c_1 \mathbf{x}^{\vec{m}_1} + \cdots + c_n \mathbf{x}^{\vec{m}_n}$$

where $\mathbf{x}^{\vec{m}_i} = x_1^{m_{1,i}} \cdots x_d^{m_{d,i}}$.

- In symbolic case : Ben-Or/Tiwari (1988).
- Basically uses a variation of Prony's method (1795).
- In numeric environment : Giesbrecht-Labahn-Lee (2006).

# Motivation : Sparse Interpolation

Given a black box representation of a multivariate polynomial $p$ :

$$\vec{x} \longrightarrow \blacksquare \longrightarrow p(\vec{x})$$

Find powers $\vec{m}_1, \ldots, \vec{m}_n$ and coefficients $c_1, \ldots, c_n$ so that

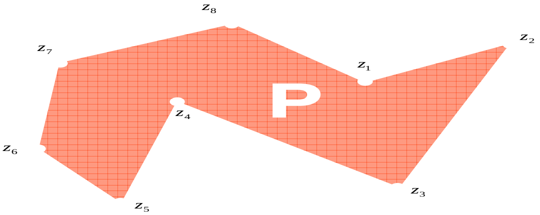
$$p(x_1, \ldots, x_d) = c_1 \mathbf{x}^{\vec{m}_1} + \cdots + c_n \mathbf{x}^{\vec{m}_n}$$

where $\mathbf{x}^{\vec{m}_i} = x_1^{m_{1,i}} \cdots x_d^{m_{d,i}}$.

- In symbolic case : Ben-Or/Tiwari (1988).
- Basically uses a variation of Prony's method (1795).
- In numeric environment : Giesbrecht-Labahn-Lee (2006).

# Motivation : Shape from Moments

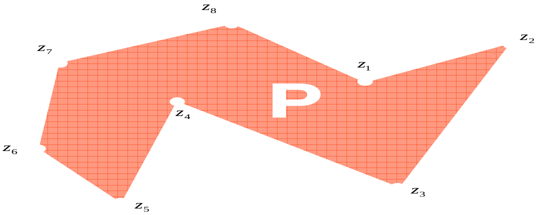
Shape reconstruction from moments (Golub, Milanfar, Varah)




- Given polygon $D$ with unknown vertices $z_1, \ldots, z_n$. Find the vertices $z_i$.
- Determine vertices from moments $h_0, \ldots, h_{2n-1}$
- Solve via generalized Hankel eigenvalue problem
  - Obtain better numerical behaviour (using QZ algorithm).

# Motivation : Shape from Moments

Shape reconstruction from moments (Golub, Milanfar, Varah)




- Given polygon $D$ with unknown vertices $z_1, \ldots, z_n$. Find the vertices $z_i$.
- Determine vertices from moments $h_0, \ldots, h_{2n-1}$
- Solve via generalized Hankel eigenvalue problem
  - Obtain better numerical behaviour (using QZ algorithm).

# Examples

**Shape from Moments** : Make use of integration formula (Davis)

$$\int\int_D f''(z)\,dxdy \quad = \quad c_1 f(z_1) + \cdots + c_n f(z_n)$$

$$\text{Set : } h_k \quad = \quad k(k-1)\int\int_D z^{k-2} dxdy = c_1 z_1^k + \cdots c_n z_n^k$$

Vertices are then $\lambda_i$'s.

**Sparse interpolation** : $h_k = p(\omega_1^k, \ldots, \omega_d^k)$,

- Over integers : $\omega_\ell = p_\ell$, prime.
- Over complex : $\omega_\ell = e^{\frac{2\pi i r_\ell}{p_\ell}}$ ($p_\ell$ relatively prime).
- Powers for each term determined after Prony (Division or CRT)

# Examples

**Shape from Moments** : Make use of integration formula (Davis)

$$\int\int_D f''(z)\, dxdy = c_1 f(z_1) + \cdots + c_n f(z_n)$$

$$\text{Set : } h_k = k(k-1)\int\int_D z^{k-2} dxdy = c_1 z_1^k + \cdots c_n z_n^k$$

Vertices are then $\lambda_i$'s.

**Sparse interpolation** : $h_k = p(\omega_1^k, \ldots, \omega_d^k)$,

- Over integers : $\omega_\ell = p_\ell$, prime.
- Over complex : $\omega_\ell = e^{\frac{2\pi i r_\ell}{p_\ell}}$ ($p_\ell$ relatively prime).
- Powers for each term determined after Prony (Division or CRT)

# Prony's Method?

Given a sequence $h_0, h_1, \ldots, h_{2n-1}$ of $2n$ values, with

$$\begin{array}{lllllllll} h_0 & = & c_1 & + & c_2 & + & \cdots & + & c_n \\ h_1 & = & c_1\lambda_1 & + & c_2\lambda_2 & + & \cdots & + & c_n\lambda_n \\ h_2 & = & c_1{\lambda_1}^2 & + & c_2{\lambda_2}^2 & + & \cdots & + & c_n{\lambda_n}^2 \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ h_{2n-1} & = & c_1{\lambda_1}^{2n-1} & + & c_2{\lambda_2}^{2n-1} & + & \cdots & + & c_n{\lambda_n}^{2n-1} \end{array}$$

- $\lambda_i$'s are roots of $a(x) = a_0 + a_1 x + \cdots + a_n x^n$. Here coefficients $a_i$ are determined by solving

$$a_0 h_i + a_1 h_{i+1} + \cdots + a_n h_{i+n-1} = -h_{i+n}\ .$$

- $c_i$'s solved via Vandermonde system

# Our Problem

**Alternately :** A generalized Hankel eigenvalue problem.

**Given :** $h_0, h_1, ..., h_{2n-1}$, solve $\tilde{H}_n \vec{y} = \lambda H_n \vec{y}$ , i.e.

$$\begin{bmatrix} h_1 & h_2 & \cdots & h_n \\ h_2 & h_3 & \cdots & h_{n+1} \\ \vdots & \vdots & & \vdots \\ h_n & h_{n+1} & \cdots & h_{2n-1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \lambda \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n-2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

for eigenvalues $\lambda$ and eigenvectors $\vec{y} = [y_1, y_2, \ldots, y_n]^T$.

# The General Problem

**Given :** $h_0, h_1, ..., h_{2n-1}$ (moments),

**Find :** $c_1, \ldots, c_n$ and $\lambda_1, \ldots, \lambda_n$ such that

$$\begin{array}{lclclclcl}
h_0 & = & c_1 & + & c_2 & + & \cdots & + & c_n \\
h_1 & = & c_1\lambda_1 & + & c_2\lambda_2 & + & \cdots & + & c_n\lambda_n \\
\vdots & & \vdots & & \vdots & & & & \vdots \\
h_{2n-1} & = & c_1{\lambda_1}^{2n-1} & + & c_2{\lambda_2}^{2n-1} & + & \cdots & + & c_n{\lambda_n}^{2n-1}
\end{array}$$

# Important Decomposition

We make heavy use of the decomposition:

$$\begin{bmatrix} h_0 & \cdots & h_{n-1} \\ h_1 & \cdots & h_n \\ \vdots & & \vdots \\ h_{n-1} & \cdots & h_{2n-2} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & \lambda_1^{n-1} \\ 1 & \cdots & \lambda_2^{n-1} \\ \vdots & & \vdots \\ 1 & \cdots & \lambda_n^{n-1} \end{bmatrix}^T \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{bmatrix} \begin{bmatrix} 1 & \cdots & \lambda_1^{n-1} \\ 1 & \cdots & \lambda_2^{n-1} \\ \vdots & & \vdots \\ 1 & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

$$H_n = V_n^T \quad diag(c_1, \ldots, c_n) \quad V_n$$

$$= W_n^T \quad W_n \tag{1}$$

Here $W_n = diag(\sqrt{c_1}, \ldots, \sqrt{c_n}) V_n$ is a scaled Vandermonde matrix. Similarly,

$$\widetilde{H}_n = V_n^T \quad diag(\lambda_1 c_1, \ldots, \lambda_n c_n) \quad V_n$$

# Issue

If $H_n$ has some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,n} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,n} \\ \vdots & \vdots & & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,n} \end{bmatrix}$$

i.e. unstructured error analysis

If $h_i$ have some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_0 & e_1 & \cdots & e_{n-1} \\ e_1 & e_2 & \cdots & e_n \\ \vdots & \vdots & & \vdots \\ e_{n-1} & e_n & \cdots & e_{2n} \end{bmatrix}$$

i.e. structured error analysis

# Issue

If $H_n$ has some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,n} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,n} \\ \vdots & \vdots & & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,n} \end{bmatrix}$$

i.e. unstructured error analysis

If $h_i$ have some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_0 & e_1 & \cdots & e_{n-1} \\ e_1 & e_2 & \cdots & e_n \\ \vdots & \vdots & & \vdots \\ e_{n-1} & e_n & \cdots & e_{2n} \end{bmatrix}$$

i.e. structured error analysis

# Issue

If $H_n$ has some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,n} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,n} \\ \vdots & \vdots & & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,n} \end{bmatrix}$$

i.e. unstructured error analysis

If $h_i$ have some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_0 & e_1 & \cdots & e_{n-1} \\ e_1 & e_2 & \cdots & e_n \\ \vdots & \vdots & & \vdots \\ e_{n-1} & e_n & \cdots & e_{2n} \end{bmatrix}$$

i.e. structured error analysis

# Issue

If $H_n$ has some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,n} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,n} \\ \vdots & \vdots & & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,n} \end{bmatrix}$$

i.e. unstructured error analysis

If $h_i$ have some noise then

$$H_n + E_n = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n} \end{bmatrix} + \begin{bmatrix} e_0 & e_1 & \cdots & e_{n-1} \\ e_1 & e_2 & \cdots & e_n \\ \vdots & \vdots & & \vdots \\ e_{n-1} & e_n & \cdots & e_{2n} \end{bmatrix}$$

i.e. structured error analysis

# Unstructured Error (GMV) : (without details)

Do usual first order error arithmetic.

Learn that unstructured errors depend on

$$(|\lambda_j| + 1) \, \|W_n^{-1} e_j\|^2 \, .$$

# Unstructured Error (GMV) : (with details)

$$[\widetilde{H}_n+\epsilon\widetilde{E_n}]y^R(\epsilon) = \lambda(\epsilon)[H_n+\epsilon E_n]y^R(\epsilon), \;\; y^L(\epsilon)[\widetilde{H}_n+\epsilon\widetilde{E_n}] = \lambda(\epsilon)y^L(\epsilon)[H_n+\epsilon E_n],$$

where $\epsilon > 0$ is small, and $\|\tilde{E}_n\| \leq 1$, $\|E_n\| \leq 1$.

$$\begin{aligned} \frac{d\lambda_j}{d\epsilon}(0) &= \frac{y^L(0)\,[\tilde{E}_n - \lambda_j E_n]\,y^R(0)}{y^L(0)\,H_n\,y^R(0)} \\ &= \frac{(V_n^{-1}e_j)^t\,[\widetilde{E_n} - \lambda_j E_n]\,V_n^{-1}e_j}{(V_n^{-1}e_j)^t\,H_n\,V_n^{-1}e_j} \\ &= \frac{(V_n^{-1}e_j)^t\,[\widetilde{E_n} - \lambda_j E_n]\,V_n^{-1}e_j}{c_j} \\ &= (W_n^{-1}e_j)^t\,[\widetilde{E_n} - \lambda_j E_n]\,W_n^{-1}e_j \leq (|\lambda_j| + 1)\,\|W_n^{-1}e_j\|^2. \end{aligned}$$

Errors depend on $(|\lambda_j| + 1)\,\|W_n^{-1}e_j\|^2$.

# Structured Error

**Theorem**

*Suppose the GEP has distinct possibly complex eigenvalues $\lambda_1, ..., \lambda_n$. Then for all $j = 1, ..., n$ we have*

$$\left\|\left(\frac{\partial \lambda_j}{\partial h_0}, \ldots, \frac{\partial \lambda_j}{\partial h_{2n-1}}\right)\right\| = \eta_{j,n} \, (|\lambda_j| + 1) \, \|W_n^{-1} e_j\|^2,$$

*where :* $\frac{1}{2n} \leq \eta_{j,n} \leq \sqrt{2n}$.

Errors depend on $(|\lambda_j| + 1) \, \|W_n^{-1} e_j\|^2$ (again).

# Structured Error

**Theorem**

*Suppose the GEP has distinct possibly complex eigenvalues $\lambda_1, ..., \lambda_n$. Then for all $j = 1, ..., n$ we have*

$$\|(\frac{\partial \lambda_j}{\partial h_0}, \ldots, \frac{\partial \lambda_j}{\partial h_{2n-1}})\| = \eta_{j,n} \, (|\lambda_j| + 1) \, \|W_n^{-1} e_j\|^2,$$

*where :* $\frac{1}{2n} \leq \eta_{j,n} \leq \sqrt{2n}$.

Errors depend on $(|\lambda_j| + 1) \, \|W_n^{-1} e_j\|^2$ (again).

# Some hints on proof:

- Inverse of Vandermonde matrices expressed via polynomials

$$V_n^{-1}\vec{e}_j = \vec{\ell}_j \quad \text{where } \ell_j(z) = \prod_{k=1,\dots,n,k\neq j} \frac{z-\lambda_k}{\lambda_j-\lambda_k}$$

- If $p(z) = \ell_j(z)$ and $q(z) = (z-\lambda_j)\ell_j(z)$ then

$$\frac{\partial\lambda_j}{\partial h_k} = \frac{(V_n^{-1}e_j)^t[\tilde{E}_n - \lambda_j E_n]V_n^{-1}e_j}{c_j} = \frac{1}{c_j}[p_{k-1} - \lambda_j p_k] = \frac{q_k}{c_j}.$$

- Need to show

$$\frac{|\lambda_j|+1}{2n}\|\vec{\ell}_j\|^2 \leq \|\vec{q}\| \leq \sqrt{2n}(|\lambda_j|+1)\,\|\vec{\ell}_j\|^2.$$

- $B\vec{p} = \vec{q}$, with $B$ being of size $(2n)\times(2n-1)$ with easily computed singular values.

# Sensitivity of GEP

Sensitivity of GEP problem related to :

- Conditioning of Hankel matrices : $H_n$ and $\widetilde{H}_n$
- Conditioning of row scaled Vandermonde matrices : $W_n$
- How $\lambda_1, \ldots, \lambda_n$ are distributed in a simply connected compact subset of $\mathbb{C}$.

Main result : If $E$ contains the generalized eigenvals, then (usually)

$$\text{Sensitivity} \geq \gamma_n(E)$$

$$n^2\,\gamma(E)^{n-1} \geq \gamma_n(E) \geq \frac{1}{\sqrt{n}}\left[\frac{\pi}{2V}\gamma(E)^{n-1} - 1 - \frac{V}{\pi}\right]$$

Here $\gamma(E) \geq 1$ measures which part of unit circle is not part of $E$.

# Tools used for Sensitivity Analysis

- Techniques from optimization and potential theory
- For a given subset $E$ of $\mathbb{C}$ the following quantities:

$$\gamma_n(E) = \inf\{\min_{D \text{ diagonal}} \|DV_n\|_F \, \|(DV_n)^{-1}\|_F : \lambda_j \in E \text{ distinct}\}$$

$$\gamma(E) = \exp\left(\max_{x \in \partial\mathbb{D}} \frac{1}{2\pi} \int_0^{2\pi} g_E(x, e^{it})\, dt\right)$$

$$\tilde{\gamma}_n(E) = \max\{\frac{\|P\|_{L_\infty(\partial\mathbb{D})}}{\|\rho^n P\|_{L_\infty(E)}} : P \text{ is a polynomial of degree } \leq n\},$$

Here $g_E(x, y)$ is the Green function for $E$ and $\rho(z) := \frac{1}{\max(|z|, 1)}$.

# Relation to conditioning of Hankel matrices

## Definition

Let

$$\rho_j := (|\lambda_j| + 1) \, \|W_n^{-1} e_j\|^2 \, (\|H_n\| + \|\widetilde{H}_n\|)$$

A generalized eigenvalue $\lambda_j$ is *ill-disposed* if $\rho_j$ is “large”.

## Theorem

*If the Hankel matrix $H_n$ is ill-conditioned then at least one of the generalized eigenvalues is ill-disposed.*

$$\rho_1 + \rho_2 + \ldots + \rho_n \geq \|H_n\| \, \|H_n^{-1}\|.$$

# Conditioning of row-scaled Vandermonde matrices

## Definition

We say that the unit disk case holds if the moments $h_k$ are generated with $|c_j| \leq 1$ and distinct $|\lambda_j| \leq 1$.

## Theorem

*Suppose that the unit disk case holds. Then*

$$\rho_1 + \rho_2 + \ldots + \rho_n \geq \frac{\|H_n\|}{n^2} \left( \|W_n\|_F \, \|W_n^{-1}\|_F \right)^2.$$

*Thus if row-scaled Vandermonde matrix $W_n$ is ill-conditioned then at least one of the gen. eigenvalues is ill-disposed.*

# Conditioning via distribution of eigenvalues

## Definition

For a given compact set $E$ in the complex plane, define

$$\gamma_n(E) = \inf\{ \min_{D \text{ diagonal}} \|DV_n\|_F \, \|(DV_n)^{-1}\|_F : \lambda_j \in E \text{ distinct}\}.$$

## Theorem

*Let $E$ be a compact set with $\lambda_1, ..., \lambda_n \in E$. Then*

$$\rho_1 + \rho_2 + ... + \rho_n \; \geq \; \gamma_n(E). \tag{2}$$

*Moreover, in the unit disk case we have that*

$$\rho_1 + \rho_2 + ... + \rho_n \; \geq \; \frac{\|H_n\|}{n^2} \gamma_n(E)^2. \tag{3}$$

# Conditioning measure $\gamma(E)$

The Green function of $E$ with pole at $y$ is given by

$$g_E(z,y) = \begin{cases} \log(|\frac{1-\overline{\phi(y)}\phi(z)}{\phi(z)-\phi(y)}|) & \text{for } z,y \in \overline{\mathbb{C}} \setminus E, \\ 0 & \text{otherwise,} \end{cases}$$

where $\phi$ the Riemann map which maps the complement of $E$ conformally onto the complement of the closed unit disk $\mathbb{D}$.

- $g_E$ is continuous in both variables and $g_E(z,y) \geq 0$ for $z,y \in \mathbb{C}$
- $g_E(z,y) > 0$ if both $z$ and $y$ lie in the complement of $E$.
- Implies

  $$\gamma(E) = \exp\left(\max_{x\in\partial\mathbb{D}} \frac{1}{2\pi}\int_0^{2\pi} g_E(x,e^{it})\,dt\right)$$

  is at least 1. Also $\gamma(E) = 1$ if and only if $\partial\mathbb{D} \subset E$.
- In some cases we have a simplified expression for $\gamma(E)$:

$$E \subset \mathbb{D}: \qquad \gamma(E) = \exp\left(\max_{x\in\partial\mathbb{D}} g_E(x,\infty)\right).$$

# Conditioning via distribution of eigenvalues

## Theorem

*For any compact set $E \subset \mathbb{C}$ which is regular with respect to the Dirichlet problem we have that*

$$\lim_{n\to\infty} \gamma_n(E)^{1/n} = \gamma(E),$$

*and $\gamma_n(E) \leq n^2\,\gamma(E)^{n-1}$ for all $n \geq 1$. If, in addition $E$ is simply connected and of bounded variation $V$, then we have that*

$$\gamma_n(E) \geq \frac{1}{\sqrt{n}}\left[\frac{\pi}{2V}\gamma(E)^{n-1} - 1 - \frac{V}{\pi}\right].$$

Note : $E$ is of *bounded variation* $V$ if, given $\beta : [0,1] \mapsto \partial E$ there exists a tangent at almost every $\beta(s)$, forming an angle $\theta(s)$ with the positive real axis, and if $\theta$ has a total variation bounded by $V$. For instance, convex sets are of bounded variation $2\pi$.

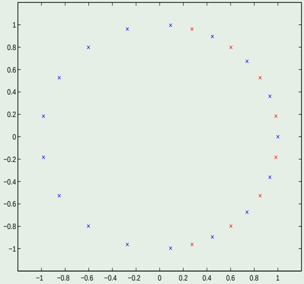
## Example (Unit Arcs)

Subarc of the unit circle $E = \{e^{it} : \theta \leq t \leq \theta + \alpha\}$ with $0 < \alpha < 2\pi$. Here $V = 2\pi + 2\alpha$. Then

$$\gamma(E) = \frac{1 + \cos(\frac{\alpha}{4})}{\sin(\frac{\alpha}{4})} = \frac{1}{\tan(\frac{\alpha}{8})}, \qquad 0 < \alpha < 2\pi.$$

## Example (Sparse Interpolation)

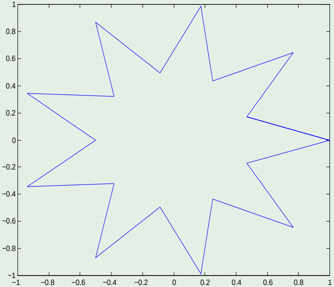
Let $E_n = \{\lambda_1, \dots, \lambda_n\}$, with $n = 6m + 1$, be the set




$$\text{Then} : \gamma_n(E_n) \geq \frac{\rho^m}{n} \text{ with } \rho = \exp\left(\frac{8\text{Catalan}}{\pi}\right) \approx 10.30 > 1$$

Hence there is still exponential growth.

## Example (Shape from Moments)




- For inner vertices, measure for ill-disposedness is exponentially increasing in $n$ unless $D$ is a regular polygon with $2n$ vertices.
- For outer vertices which do not lie on the unit circle but strictly inside the unit disk, then our measure for ill-disposedness is exponentially increasing in $n$.
- Thus one should scale $D$ in a way that it takes as much space as possible in the unit disk.