MANCHESTER 1824

The University of Manchester

The Conditioning of Linearizations of Matrix Polynomials

Nick Higham
School of Mathematics
University of Manchester

higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/

Joint work with **D. Steven Mackey** and **Françoise Tisseur**.

SIAM Annual Meeting, New Orleans, July 2005

Polynomial Eigenproblem

$$P(\lambda) = \sum_{i=0}^{m} \lambda^i A_i, \qquad A_i \in \mathbb{C}^{n\times n}, \quad A_m \neq 0.$$

P assumed **regular** ($\det P(\lambda) \not\equiv 0$).

Find scalars λ and nonzero vectors x and y satisfying $P(\lambda)x = 0$ and $y^* P(\lambda) = 0$.

- Standard eigenvalue problem: $Ax = \lambda x.$
- Generalized eigenvalue problem (GEP): $Ax = \lambda Bx.$
- Quadratic eigenvalue problem (QEP): $(\lambda^2 M + \lambda C + K)x = 0$.

Linearizations

The pencil

$$L(\lambda) = \lambda X + Y, \qquad X, Y \in \mathbb{C}^{mn \times mn}$$

is a **linearization** of $P(\lambda) = \sum_{i=0}^{m} \lambda^i A_i$ if

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(m-1)n} \end{bmatrix}$$

for some unimodular $E(\lambda)$ and $F(\lambda)$.

Standard way of solving $P(\lambda)x = 0$:

- Linearize $P(\lambda)$ into $L(\lambda) = \lambda X + Y$.
- Solve generalized eigenproblem $L(\lambda)z = 0$.

$\mathbb{L}_1$ and $\mathbb{L}_2$ Linearizations

$$\Lambda := [\lambda^{m-1}, \lambda^{m-2}, \dots, 1]^T.$$

Mackey, Mackey, Mehl & Mehrmann (2005) define

$$\mathbb{L}_1(P) = \{\, L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda),\ v \in \mathbb{C}^m \,\},$$

$$\mathbb{L}_2(P) = \{\, L(\lambda) : (\Lambda^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda),\ w \in \mathbb{C}^m \,\}.$$

They show that

- $\mathbb{L}_1$ and $\mathbb{L}_2$ are vector spaces.
- Almost all pencils in $\mathbb{L}_1$ and $\mathbb{L}_2$ are linearizations.

Quadratic case ($m = 2$): $L = \lambda X + Y \in \mathbb{L}_1(P)$ iff

$$\begin{bmatrix} v_1A_2 & v_1A_1 & v_1A_0 \\ v_2A_2 & v_2A_1 & v_2A_0 \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} + Y_{11} & Y_{12} \\ X_{21} & X_{22} + Y_{21} & Y_{22} \end{bmatrix}.$$

$\mathbb{L}_1$ and $\mathbb{L}_2$ Linearizations cont.

Recall

$$\mathbb{L}_1(P) = \{\, L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda),\ v \in \mathbb{C}^m \,\}.$$

Note

$$L(\lambda)(\Lambda \otimes x) = L(\lambda)(\Lambda \otimes I_n)(1 \otimes x) = (v \otimes P(\lambda))(1 \otimes x) = v \otimes P(\lambda)x.$$

So (x, λ) **is an ei'pair of P iff $(\Lambda \otimes x, \lambda)$ is an ei'pair of L.**

- **Right** eigenvectors of P can be recovered from **right** eigenvectors of linearizations in $\mathbb{L}_1$.
- **Left** eigenvectors of P can be recovered from **left** eigenvectors of linearizations in $\mathbb{L}_2$.

$\mathbb{DL}(P)$ Linearizations

Mackey, Mackey, Mehl & Mehrmann (2005) define

$$\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P).$$

They show that

- $L \in \mathbb{DL}(P)$ iff $w = v$ in the definitions of $\mathbb{L}_1$ and $\mathbb{L}_2$.
- $\mathbb{DL}(P)$ is a vector space of dimension m.

In this work we focus on linearizations in $\mathbb{DL}(P)$.

$\mathbb{DL}(P)$ Linearizations

Mackey, Mackey, Mehl & Mehrmann (2005) define

$$\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P).$$

They show that

- $L \in \mathbb{DL}(P)$ iff $w = v$ in the definitions of $\mathbb{L}_1$ and $\mathbb{L}_2$.
- $\mathbb{DL}(P)$ is a vector space of dimension m.

In this work we focus on linearizations in $\mathbb{DL}(P)$.

Example: For $Q(\lambda) = \lambda^2 A + \lambda B + C$, $\mathbb{DL}(Q)$ is the pencils

$$L(\lambda) = \lambda \begin{bmatrix} v_1 A & v_2 A \\ v_2 A & v_2 B - v_1 C \end{bmatrix} + \begin{bmatrix} v_1 B - v_2 A & v_1 C \\ v_1 C & v_2 C \end{bmatrix}, \quad v \in \mathbb{C}^2.$$

Key Questions

- For given λ which $L \in \mathbb{DL}(P)$ has the **smallest condition number** for λ (minimization over v).
- How does this minimal condition number compare with **condition number of P** for λ?
- How well conditioned is L for $v = e_1$ and $v = e_m$?

Condition Number for $P(\lambda)$

Simple λ, $P(\lambda)x = 0$, $y^*P(\lambda) = 0$.

$$\kappa_P(\lambda) = \lim_{\epsilon \to 0} \sup \left\{ \frac{|\Delta\lambda|}{\epsilon|\lambda|} : \big(P(\lambda + \Delta\lambda) + \Delta P(\lambda + \Delta\lambda)\big)(x + \Delta x) = 0, \right. \\ \left. \|\Delta A_i\|_2 \le \epsilon\omega_i,\ i = 0\colon m \right\}.$$

Tisseur (2000) showed

$$\kappa_P(\lambda) = \frac{\left(\sum_{i=0}^m |\lambda|^i \omega_i\right)\|y\|_2\|x\|_2}{|\lambda|\,|y^*P'(\lambda)x|}.$$

Problems

- κ_P not defined for $\lambda = 0$.
- κ_P not defined for $\lambda = \infty$.

Condition Number for Homogeneous Form

$$P(\alpha,\beta) = \sum_{i=0}^{m} \alpha^i \beta^{m-i} A_i \qquad (\lambda \equiv \alpha/\beta).$$

E'vals are pairs $(\alpha,\beta) \neq (0,0)$ s.t. $\det P(\alpha,\beta) = 0$.
With $\|A\| := \|[\omega_0^{-1}A_0,\ldots,\omega_m^{-1}A_m]\|_F$,

$$\kappa_P(\alpha,\beta) = \left(\sum_{i=0}^{m} |\alpha|^{2i}|\beta|^{2(m-i)}\omega_i^2\right)^{1/2} \frac{\|x\|_2\|y\|_2}{\left|y^*(\bar{\beta}\mathcal{D}_\alpha P - \bar{\alpha}\mathcal{D}_\beta P)|_{(\alpha,\beta)}x\right|},$$

where $\mathcal{D}_\alpha \equiv \frac{\partial}{\partial\alpha}$, $\mathcal{D}_\beta \equiv \frac{\partial}{\partial\beta}$.

Condition Number for Homogeneous Form

$$P(\alpha,\beta) = \sum_{i=0}^{m} \alpha^i \beta^{m-i} A_i \qquad (\lambda \equiv \alpha/\beta).$$

E'vals are pairs $(\alpha,\beta) \neq (0,0)$ s.t. $\det P(\alpha,\beta) = 0$.
With $\|A\| := \|[\omega_0^{-1} A_0, \ldots, \omega_m^{-1} A_m]\|_F$,

$$\kappa_P(\alpha,\beta) = \left(\sum_{i=0}^{m} |\alpha|^{2i}|\beta|^{2(m-i)}\omega_i^2\right)^{1/2} \frac{\|x\|_2\|y\|_2}{\left|y^*(\bar{\beta}\mathcal{D}_\alpha P - \bar{\alpha}\mathcal{D}_\beta P)|_{(\alpha,\beta)} x\right|},$$

where $\mathcal{D}_\alpha \equiv \frac{\partial}{\partial\alpha}$, $\mathcal{D}_\beta \equiv \frac{\partial}{\partial\beta}$.

Can show that, with suitable normalizations,

$$\left|\theta\big((\alpha,\beta),(\widetilde{\alpha},\widetilde{\beta})\big)\right| \leq \kappa_P(\alpha,\beta)\|\Delta A\| + o(\|\Delta A\|).$$

where $\Delta A \equiv (\Delta A_0, \ldots, \Delta A_m)$.

Condition Number κ_L (1)

$$\Lambda_{\alpha,\beta} = [\alpha^{m-1}, \alpha^{m-2}\beta, \ldots, \beta^{m-1}]^T.$$

$$L(\alpha,\beta)z = 0, \quad w^* L(\alpha,\beta) = 0, \qquad w = \overline{\Lambda}_{\alpha,\beta} \otimes y, \quad z = \Lambda_{\alpha,\beta} \otimes x.$$

$$\kappa_L(\alpha,\beta) = \sqrt{|\alpha|^2\omega_X^2 + |\beta|^2\omega_Y^2}\, \frac{\|w\|_2 \|z\|_2}{\left|w^*(\bar\beta\,\mathcal{D}_\alpha L - \bar\alpha\,\mathcal{D}_\beta L)|_{(\alpha,\beta)} z\right|}.$$

$$L \in \mathbb{L}_1 \Rightarrow \qquad L(\alpha,\beta)(\Lambda_{\alpha,\beta} \otimes I_n) = v \otimes P(\alpha,\beta).$$

Condition Number κ_L (1)

$$\Lambda_{\alpha,\beta} = [\alpha^{m-1}, \alpha^{m-2}\beta, \ldots, \beta^{m-1}]^T.$$

$$L(\alpha,\beta)z = 0, \quad w^*L(\alpha,\beta) = 0, \quad w = \overline{\Lambda}_{\alpha,\beta} \otimes y, \quad z = \Lambda_{\alpha,\beta} \otimes x.$$

$$\kappa_L(\alpha,\beta) = \sqrt{|\alpha|^2\omega_X^2 + |\beta|^2\omega_Y^2}\, \frac{\|w\|_2\|z\|_2}{\left|w^*(\bar{\beta}\,\mathcal{D}_\alpha L - \bar{\alpha}\,\mathcal{D}_\beta L)|_{(\alpha,\beta)}z\right|}.$$

$$L \in \mathbb{L}_1 \Rightarrow \quad L(\alpha,\beta)(\Lambda_{\alpha,\beta} \otimes I_n) = v \otimes P(\alpha,\beta).$$

Differentiate wrt α:

$$\mathcal{D}_\alpha L(\alpha,\beta)(\Lambda_{\alpha,\beta} \otimes I_n) + L(\alpha,\beta)(\mathcal{D}_\alpha\Lambda_{\alpha,\beta} \otimes I_n) = v \otimes \mathcal{D}_\alpha P(\alpha,\beta).$$

Condition Number κ_L (1)

$$\Lambda_{\alpha,\beta} = [\alpha^{m-1}, \alpha^{m-2}\beta, \ldots, \beta^{m-1}]^T.$$

$$L(\alpha,\beta)z = 0, \quad w^* L(\alpha,\beta) = 0, \quad w = \overline{\Lambda}_{\alpha,\beta} \otimes y, \quad z = \Lambda_{\alpha,\beta} \otimes x.$$

$$\kappa_L(\alpha,\beta) = \sqrt{|\alpha|^2\omega_X^2 + |\beta|^2\omega_Y^2}\, \frac{\|w\|_2\|z\|_2}{\left|w^*(\bar{\beta}\mathcal{D}_\alpha L - \bar{\alpha}\mathcal{D}_\beta L)|_{(\alpha,\beta)}z\right|}.$$

$$L \in \mathbb{L}_1 \Rightarrow \quad L(\alpha,\beta)(\Lambda_{\alpha,\beta} \otimes I_n) = v \otimes P(\alpha,\beta).$$

Differentiate wrt α:

$$\mathcal{D}_\alpha L(\alpha,\beta)(\Lambda_{\alpha,\beta} \otimes I_n) + L(\alpha,\beta)(\mathcal{D}_\alpha \Lambda_{\alpha,\beta} \otimes I_n) = v \otimes \mathcal{D}_\alpha P(\alpha,\beta).$$

Multiplying on left by w^* and on right by $1 \otimes x$:

$$\begin{aligned} w^*\mathcal{D}_\alpha L(\alpha,\beta)z &= \Lambda_{\alpha,\beta}^T v \otimes y^*\mathcal{D}_\alpha P(\alpha,\beta)x \\ &= \Lambda_{\alpha,\beta}^T v \cdot y^*\mathcal{D}_\alpha P(\alpha,\beta)x. \end{aligned}$$

Condition Number κ_L (2)

Theorem 1 *For any pencil* $L(\alpha,\beta) = \alpha X + \beta Y \in \mathbb{DL}(P)$,

$$\kappa_L(\alpha,\beta) = \frac{\sqrt{|\alpha|^2\omega_X^2 + |\beta|^2\omega_Y^2}\, \|\Lambda_{\alpha,\beta}\|_2^2 \;\; \|x\|_2\|y\|_2}{|\mathsf{p}(\alpha,\beta;v)|\, \left|y^*(\bar{\beta}\mathcal{D}_\alpha P - \bar{\alpha}\mathcal{D}_\beta P)|_{(\alpha,\beta)}x\right|},$$

where $\mathsf{p}(\alpha,\beta;v) = \Lambda_{\alpha,\beta}^T v = \sum_{i=1}^m v_i\alpha^{m-i}\beta^{i-1}$.

Condition Number κ_L (2)

Theorem 1 *For any pencil* $L(\alpha, \beta) = \alpha X + \beta Y \in \mathbb{DL}(P)$,

$$\kappa_L(\alpha, \beta) = \frac{\sqrt{|\alpha|^2 \omega_X^2 + |\beta|^2 \omega_Y^2}\, \|\Lambda_{\alpha,\beta}\|_2^2 \;\; \|x\|_2 \|y\|_2}{|\mathsf{p}(\alpha, \beta; v)| \left| y^* (\bar{\beta} \mathcal{D}_\alpha P - \bar{\alpha} \mathcal{D}_\beta P)|_{(\alpha,\beta)} x \right|},$$

where $\mathsf{p}(\alpha, \beta; v) = \Lambda_{\alpha,\beta}^T v = \sum_{i=1}^m v_i \alpha^{m-i} \beta^{i-1}$.

- **Second term** in denominator is same as in $\kappa_P(\alpha, \beta)$ expression.

Condition Number κ_L (2)

Theorem 1 *For any pencil* $L(\alpha,\beta) = \alpha X + \beta Y \in \mathbb{DL}(P)$,

$$\kappa_L(\alpha,\beta) = \frac{\sqrt{|\alpha|^2\omega_X^2 + |\beta|^2\omega_Y^2}\,\|\Lambda_{\alpha,\beta}\|_2^2\ \ \|x\|_2\|y\|_2}{|\mathsf{p}(\alpha,\beta;v)|\,\left|y^*(\bar{\beta}\mathcal{D}_\alpha P - \bar{\alpha}\mathcal{D}_\beta P)|_{(\alpha,\beta)}x\right|},$$

where $\mathsf{p}(\alpha,\beta;v) = \Lambda_{\alpha,\beta}^T v = \sum_{i=1}^m v_i\alpha^{m-i}\beta^{i-1}$.

- **Second term** in denominator is same as in $\kappa_P(\alpha,\beta)$ expression.
- $\kappa_L(\alpha,\beta)$ is finite iff (α,β) is not a zero of $\mathsf{p}(\alpha,\beta;v)$. **Cf. M^4 paper**: $L(\lambda)$ is a linearization for $P(\lambda)$ iff no eigenvalue of P is a root of $\mathsf{p}(\lambda;v)$.

Minimizing the Condition Number κ_L (1)

Pencil $L(\lambda) \in \mathbb{DL}(P)$ defined by vector $v \in \mathbb{C}^m$.
Aim: minimize $\kappa_L(\alpha, \beta)$ over all v (with $\|v\|_2 = 1$ wlog).
Take

$$\omega_X = \|X\|_2, \quad \omega_Y = \|Y\|_2.$$

Can bound **numerator** of κ_L above and below.
To maximize **denominator** of κ_L, **maximize p$(\alpha, \beta; v)$**.
Cauchy–Schwarz gives

$$v_* = \frac{\overline{\Lambda}_{\alpha,\beta}}{\|\Lambda_{\alpha,\beta}\|_2}.$$

Important special cases:

$$\begin{aligned} (\alpha, \beta) &= (1, 0), \quad \lambda = \infty \quad \Rightarrow \quad v_* = e_1, \\ (\alpha, \beta) &= (0, 1), \quad \lambda = 0 \quad \Rightarrow \quad v_* = e_m. \end{aligned}$$

Minimizing the Condition Number κ_L (2)

Theorem 2 *Consider pencils* $L = \alpha X + \beta Y \in \mathbb{DL}(P)$*. Set weights* $\omega_X = \|X\|_2$, $\omega_Y = \|Y\|_2$*. Then*

$$\kappa_L(\alpha, \beta; e_1) \le \rho m^{3/2} \inf_v \kappa_L(\alpha, \beta, v) \text{ if } A_0 \text{ nonsing, } |\alpha| \ge |\beta|,$$

$$\kappa_L(\alpha, \beta; e_m) \le \rho m^{3/2} \inf_v \kappa_L(\alpha, \beta, v) \text{ if } A_m \text{ nonsing, } |\alpha| \le |\beta|,$$

where $$\rho = \frac{\max_i \|A_i\|_2}{\min(\|A_0\|_2, \|A_m\|_2)}.$$

- For $\rho = O(1)$, **one of** $v = e_1$ **and** $v = e_m$ gives **near optimal** κ_L for λ.
- **Wrong choice** of $v = e_1$ or $v = e_m$ can be **disastrous**: $\kappa_L(0, \beta, e_1) = \infty$, $\kappa_L(\alpha, 0, e_m) = \infty$.

Optimal κ_L Versus κ_P

Theorem 3 *Let* (α, β) *be a simple eigenvalue of* P*. Then*

$$\frac{1}{\rho} \le \frac{\inf_v \kappa_L(\alpha,\beta;v)}{\kappa_P(\alpha,\beta)} \le m^2 \rho.$$

If $\rho = O(1)$:

- Best conditioned $L \in \mathbb{DL}(P)$ for a given λ is about as well conditioned as P itself for λ.

  Despite perturbations to L not respecting the block structure of X and Y!

- Combined with Thm. 2:
  one of pencils with $v = e_1$ and $v = e_m$ is about as well conditioned as P itself for λ.

Quadratic Case: $Q(\lambda) = \lambda^2 A + \lambda B + C$

Write $\quad a = \|A\|_2, \quad b = \|B\|_2, \quad c = \|C\|_2.$

For optimality of $\kappa_L(\alpha, \beta)$ for $v = e_1$, $v = e_2$ need $\rho = O(1)$, i.e.

$$b \lesssim \max(a, c) \quad \text{and} \quad a \approx c.$$

With the scaling : $\lambda^2 A + \lambda B + C \to \mu^2(\gamma^2 A) + \mu(\gamma B) + C$ $(\lambda = \mu\gamma)$, $\gamma = \sqrt{c/a}$ [Fan, Lin & Van Dooren, 2004], it suffices that

$$b \lesssim \sqrt{ac}. \qquad (*)$$

$(*)$ is true for elliptic QEPs.

Companion Linearizations (1)

$C_i(\lambda) = \lambda X + Y_i$ with $X = \mathrm{diag}(A_m, I_n, \dots, I_n)$,

$$Y_1 = \begin{bmatrix} A_{m-1} & A_{m-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -I_n & 0 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} A_{m-1} & -I_n & \dots & 0 \\ A_{m-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & 0 & \dots & 0 \end{bmatrix}.$$

- $C_1 \in \mathbb{L}_1(P)$ ($v = e_1$) and $C_2 \in \mathbb{L}_2(P)$ ($w = e_1$), but $C_1, C_2 \notin \mathbb{DL}(P)$.
- C_1 and C_2 *always* linearizations.
- Know left e'vec of C_1 in terms of left e'vec of P.
- Have formula for $\kappa_{C_1}(\lambda)$.

Companion Linearizations (2)

Find that κ_{C_1}/κ_P and $\kappa_{C_1}/\kappa_L(\lambda; v_*)$ depend on

- ratios $\|w\|_2/\|y\|_2$ of norms of left ei'vecs of C_1 and P,
- rational functions of the $\|A_i\|_2$s and λ.

Conclude that

- If $\|A_i\| \approx 1$, $i = 0\colon m$ then $\kappa_{C_1} \approx \kappa_P$ and $\kappa_{C_1} \approx \kappa_L(\lambda; v_*)$.
- If $\|w\|_2/\|y\|_2 \gg 1$ or if $\|A_i\|_2 \ll 1$, $i = 0\colon m$, then $\kappa_{C_1} \gg \kappa_L(\lambda; v_*)$ is possible.

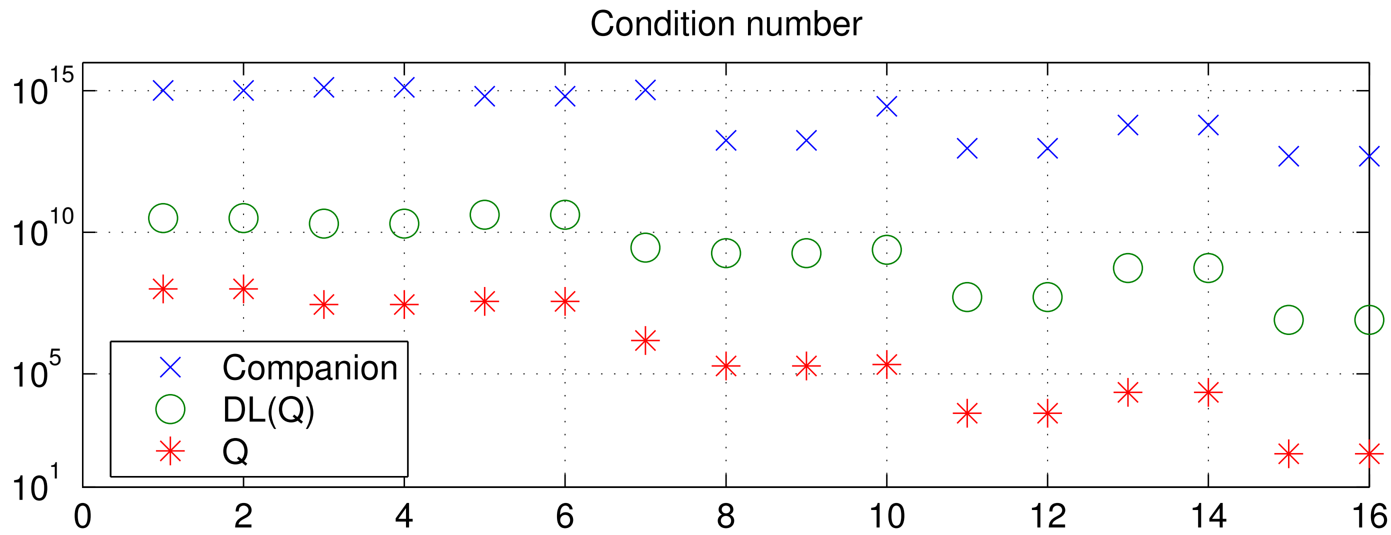
Same comments apply to C_2 since

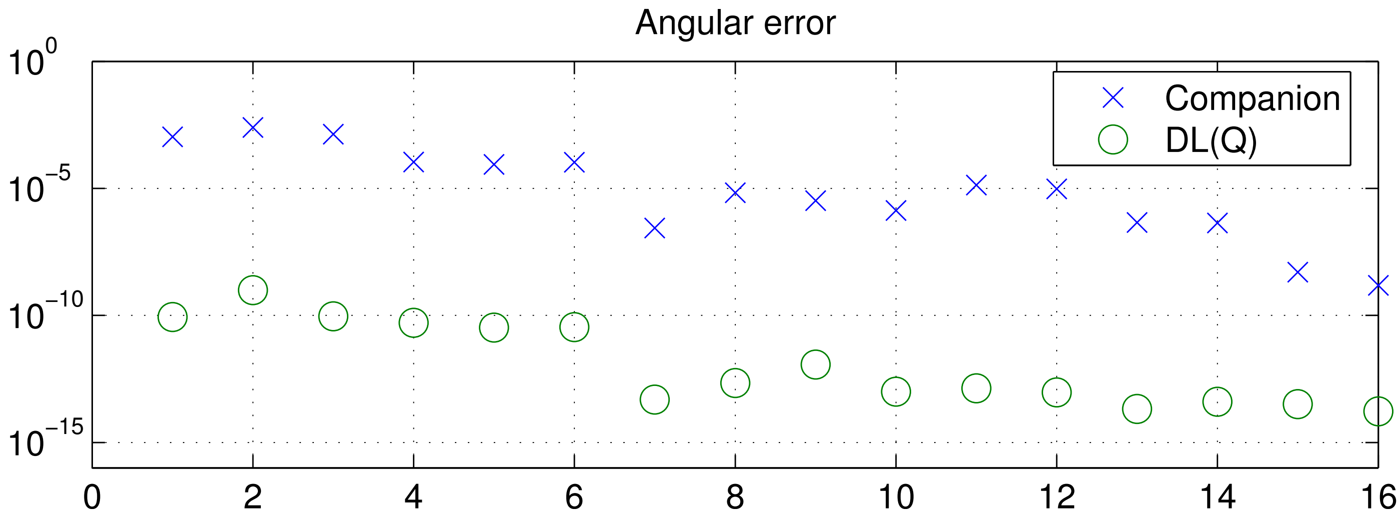
$$\kappa_{C_2(P)}(\lambda) = \kappa_{C_1(P^T)}(\lambda).$$

Example 1: Nuclear Power Plant (1)

$n = 8$; $\|A\|_2 = 2.3 \times 10^8$, $\|B\|_2 = 4.3 \times 10^{10}$, $\|C\|_2 = 1.7 \times 10^{13}$.

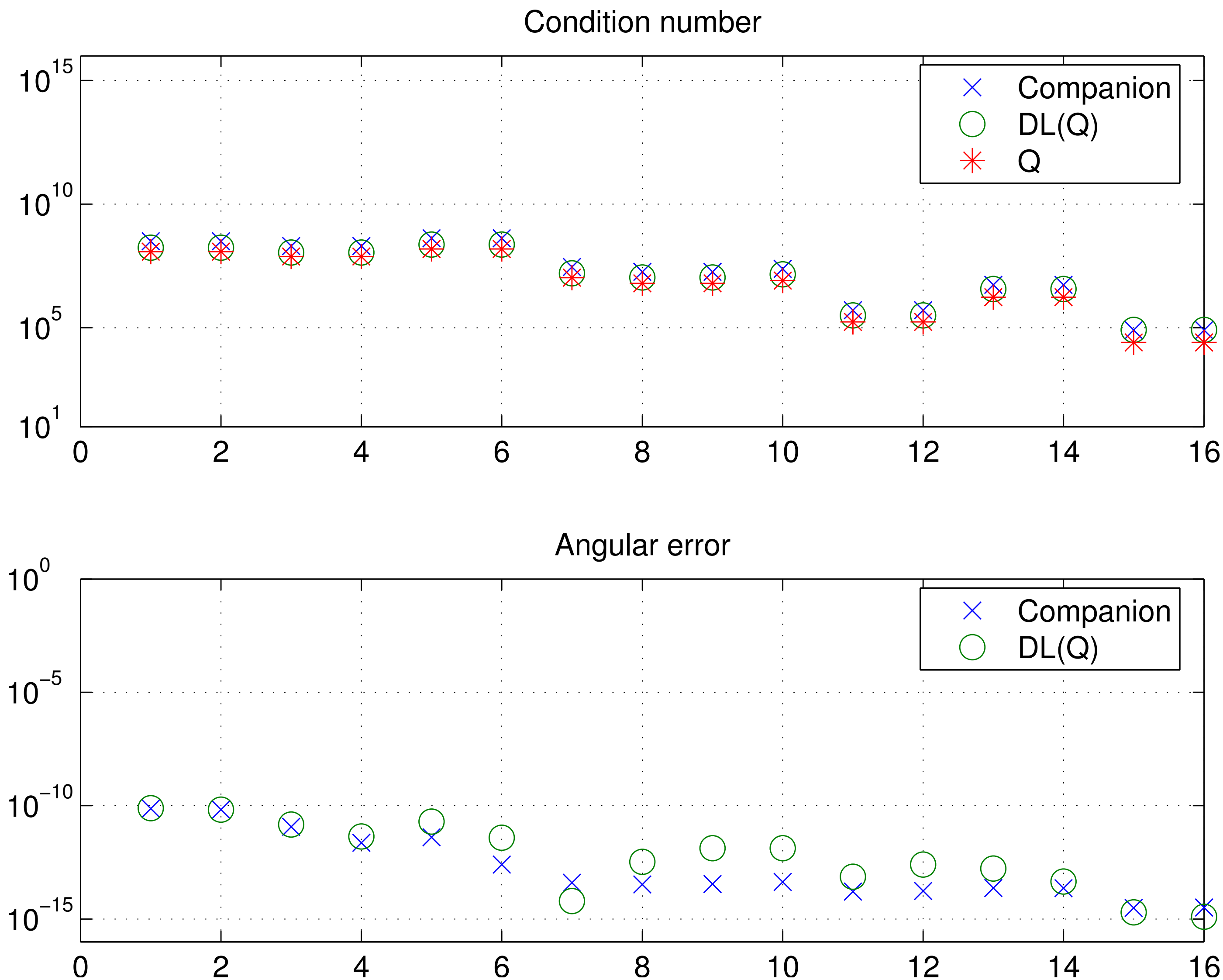
$|\lambda| \in (17, 362)$; Unscaled; $v = e_1$, $\rho = 7 \times 10^4$.





Example 1: Nuclear Power Plant (2)

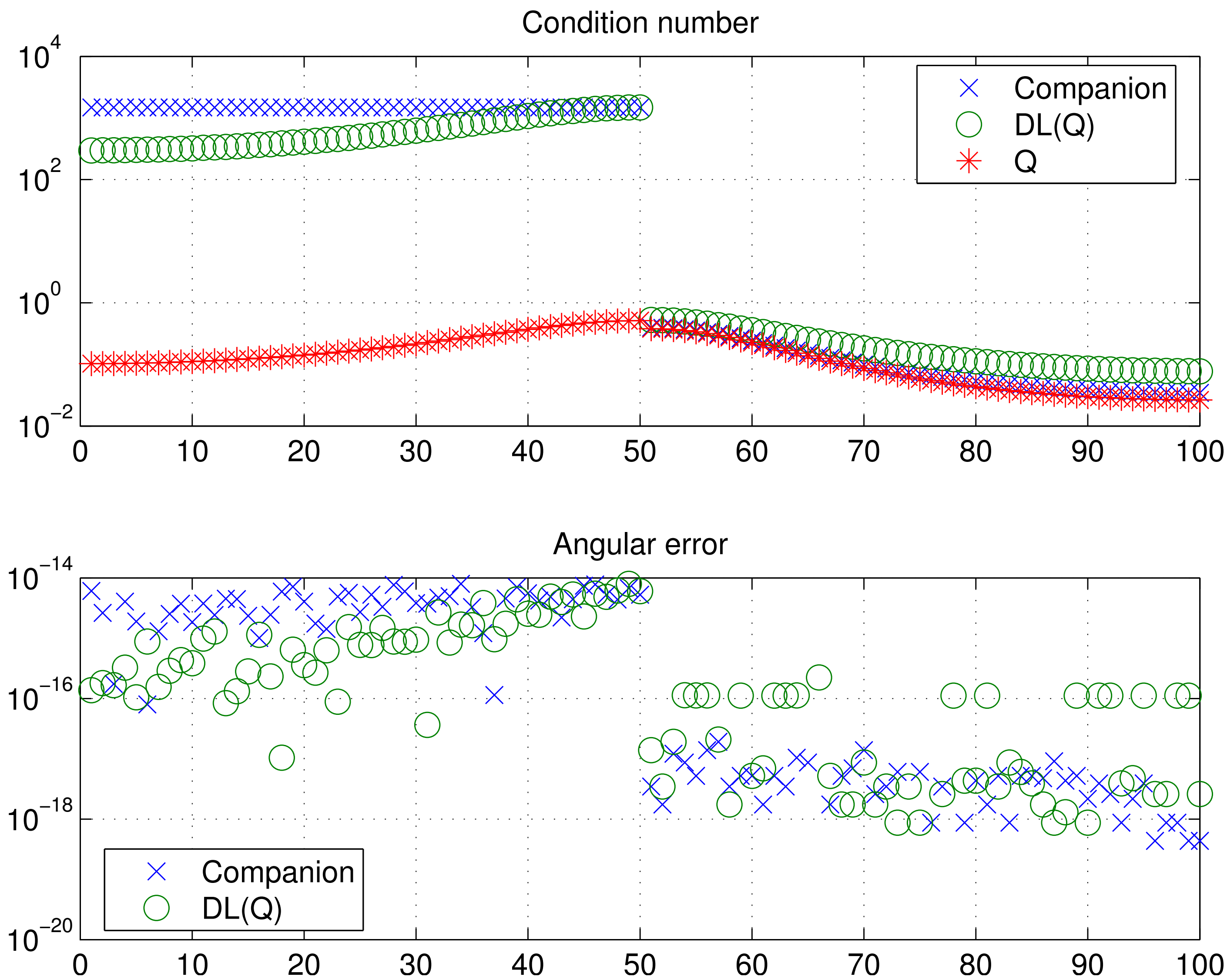
Scaled; $v = e_2$, $\rho = 1$.



Example 2: Damped Mass-Spring (1)

$n = 50$; 50 $|\lambda| \in (-320, -6.4)$, 50 $\lambda \approx -1.5 \times 10^{-2}$.

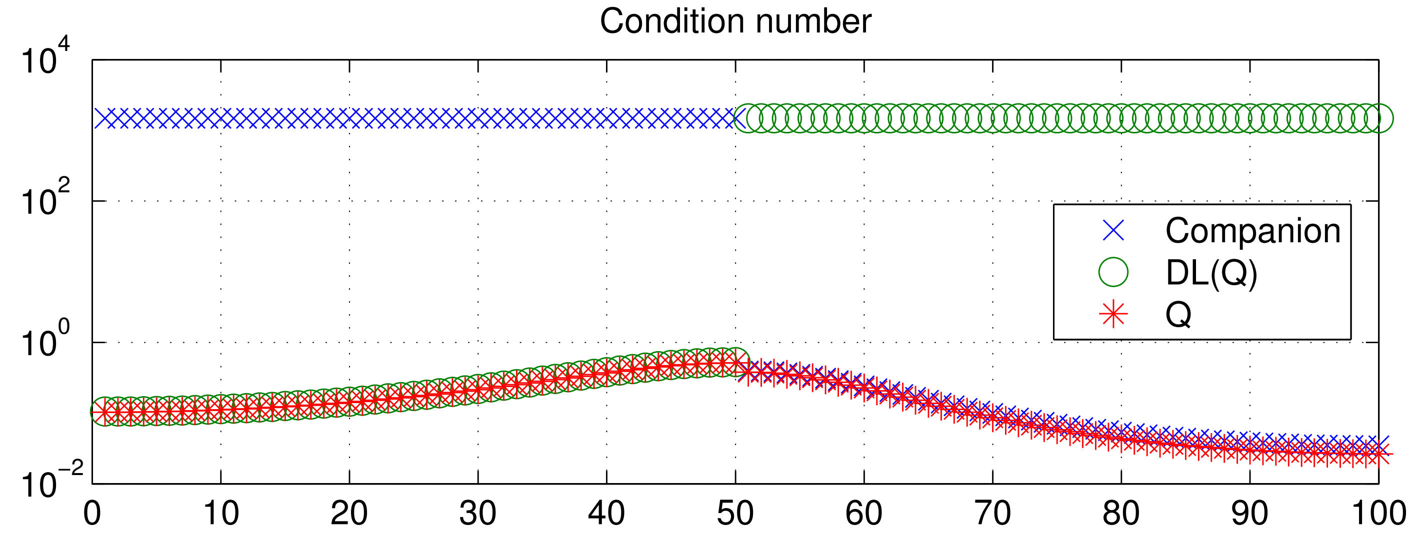
Unscaled; $v = e_1$, $\rho = 320$.

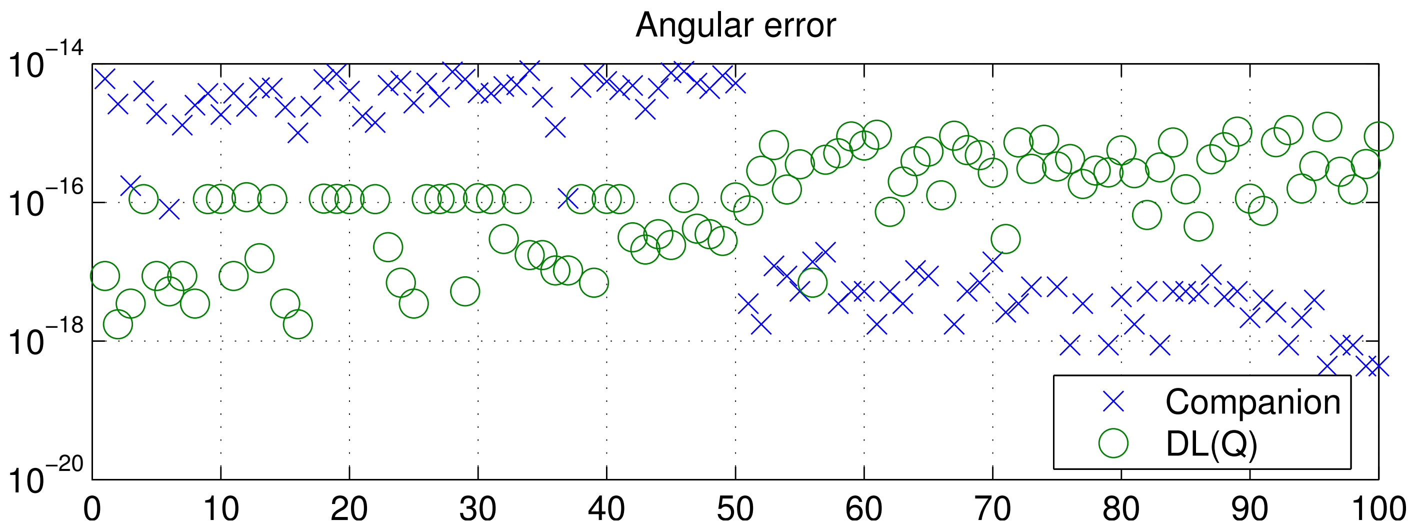


Example 2: Damped Mass-Spring (2)

$n = 50$; 50 $|\lambda| \in (-320, -6.4)$, 50 $\lambda \approx -1.5 \times 10^{-2}$.

Unscaled; $v = e_2$, $\rho = 320$.





Conclusions

$$P(\lambda) = \sum_{i=0}^{m} \lambda^i A_i, \qquad L(\lambda) = \lambda X + Y \in \mathbb{DL}(P).$$

Analyzed ei'val conditioning of linearizations in $\mathbb{DL}(P)$.

- ★ Showed that if $\max_i \|A_i\|_2 \approx \min(\|A_0\|_2, \|A_m\|_2)$ then L for $v = e_1$ and $v = e_m$ near **optimally conditioned** for $|\lambda| \geq 1$ and $|\lambda| \leq 1$, resp. *and* as well conditioned as P.
- ★ For quadratics, assumption on coeffs weakened to $\|B\|_2 \lesssim \sqrt{\|A\|_2 \|C\|_2}$ by use of scaling.
- ★ Companion linearizations can be poorly conditioned.

Justification for solving $P(\lambda)x = 0$ by linearization.