Solving Polynomial Eigenproblems by Linearization

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**

13th ACME Annual Conference, Sheffield, March 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 (SEP): $Ax = \lambda x$.
- Generalized eigenvalue problem (GEP): $Ax = \lambda Bx$.
- Quadratic eigenvalue problem (QEP): $(\lambda^2 M + \lambda C + K)x = 0$.

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

Have very good ways of solving SEP, GEP but not QEP.

Applications of QEP

From, e.g., vibration analysis of structural systems,

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t).$$

Leads to QEP

$$(\lambda^2 M + \lambda C + K)x = 0.$$

Damping term C is arbitrary.

More applications

- Acoustic structural coupled systems
- Fluid mechanics
- MIMO systems in control theory
- Signal processing (time series forecasting)
- Constrained least squares problems

Spectrum of P

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

Recall p assumed regular. Note that

$$\begin{aligned}\det(P(\lambda)) &= \lambda^{mn}\det(A_m) + \alpha_{mn-1}\lambda^{mn-1} + \cdots + \det(A_0) \\ &= \alpha_r\lambda^r + \cdots + \alpha_1\lambda + \alpha_0, \qquad \alpha_r \neq 0.\end{aligned}$$

- P has r finite eigenvalues: the roots of $\det(P(\lambda)) = 0$.
- P has $mn - r$ infinite eigenvalues.

Spectrum of P

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

Recall p assumed regular. Note that

$$\begin{aligned}\det(P(\lambda)) &= \lambda^{mn} \det(A_m) + \alpha_{mn-1}\lambda^{mn-1} + \cdots + \det(A_0)\\ &= \alpha_r \lambda^r + \cdots + \alpha_1\lambda + \alpha_0, \qquad \alpha_r \neq 0.\end{aligned}$$

- P has r finite eigenvalues: the roots of $\det(P(\lambda)) = 0$.
- P has $mn - r$ infinite eigenvalues.
- A_m singular implies $\lambda = \infty$ is an eigenvalue.
- A_0 singular implies $\lambda = 0$ is an eigenvalue.
- $\lambda = \infty$ for $P(\lambda)$ corr. to $\lambda = 0$ for $\lambda^{mn} P(1/\lambda)$.

Polynomial Zeros

The roots of

$$p(\lambda) = \alpha_m \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_0, \qquad a_m \neq 0,$$

are the eigenvalues of the **companion matrix**

$$C = \begin{bmatrix} -a_{m-1}/a_m & -a_{m-2}/a_{m-1} & \dots & \dots & -a_0/a_m \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & 0 & 0 \\ 0 & \dots & \dots & 1 & 0 \end{bmatrix}.$$

MATLAB's **`roots`** computes polynomial roots by applying **`eig`** to C.

Nonlinear polynomial $\Rightarrow$ linear eigenproblem.

Linearization

Reduction of 2nd order DE to 1st order:

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t), \qquad q, f \in \mathbb{C}^n.$$

Define $p_1 = q$, $p_2 = p_1'$. Then

$$\begin{aligned} p_1' &= p_2, \\ Mp_2' + Cp_2 + Kp_1 &= f. \end{aligned}$$

Linearization

Reduction of 2nd order DE to 1st order:

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t), \qquad q, f \in \mathbb{C}^n.$$

Define $p_1 = q$, $p_2 = p_1'$. Then

$$\begin{aligned} p_1' &= p_2, \\ Mp_2' + Cp_2 + Kp_1 &= f. \end{aligned}$$

Reduction of QEP to GEP:

$$(\lambda^2 M + \lambda C + K)x = 0.$$

Define $y_1 = x$, $y_2 = \lambda x$. Then

$$\left(\lambda \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} + \begin{bmatrix} 0 & -I \\ C & K \end{bmatrix} \right) y = 0.$$

Linearization Definition

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 $F(\lambda)$.

Example: Companion form linearization

$$E(\lambda)\left(\lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} C & K \\ -I & 0 \end{bmatrix}\right) F(\lambda) = \begin{bmatrix} \lambda^2 M + \lambda C + K & 0 \\ 0 & I \end{bmatrix}.$$

Some Linearizations of $\lambda^2 M + \lambda C + K$

L1 : $\lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} C & K \\ -I & 0 \end{bmatrix}$ **first companion**

L2 : $\lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} C & -I \\ K & 0 \end{bmatrix}$ **second companion**

L3 : $\lambda \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} + \begin{bmatrix} C & K \\ K & 0 \end{bmatrix}$ $\mathbb{DL}(Q)$

L4 : $\lambda \begin{bmatrix} 0 & M \\ M & C \end{bmatrix} + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}$ $\mathbb{DL}(Q)$

Choice of Linearization?

- ★ Which best preserves **structure** (symmetry, Hamiltonian, . . .) in Q?
- ★ Which gives the **most accurate** computed eigenvalues?
- ★ Which is the **best conditioned** for a particular eigenvalue?
- ★ How does conditioning of L compare with that of Q?

$\mathbb{DL}(P)$ Linearizations

Mackey, Mackey, Mehl & Mehrmann (2005) identify an **m-dimensional vector space of linearizations for $P(\lambda)$** s.t.

- right (left) e'vecs of P can be recovered from right (left) e'vecs of L.

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

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

- Right e'vecs of L are $\begin{bmatrix} \lambda x \\ x \end{bmatrix}$, where $x =$ right e'vec of Q.

(Technicality: L is a linearization if no e'val of Q equals $-v_2/v_1$.)

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

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

Minimizing the Condition Number κ_L

Theorem 1 *Consider pencils* $L = \lambda X + Y \in \mathbb{DL}(P)$. *Set relative weights* $\omega_i = \|A_i\|_2$. *Then*

$$\kappa_L(\lambda; e_1) \le \rho m^{3/2} \inf_v \kappa_L(\lambda, v) \quad \text{if } A_0 \text{ nonsing,} \quad |\lambda| \ge 1,$$

$$\kappa_L(\lambda; e_m) \le \rho m^{3/2} \inf_v \kappa_L(\lambda, v) \quad \text{if } A_m \text{ nonsing,} \quad |\lambda| \le 1,$$

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**: $\kappa_L(0, e_1) = \infty$, $\kappa_L(\infty, e_m) = \infty$.

QEP Case

For $Q(\lambda) = \lambda^2 A + \lambda B + C$,

$$v = e_1 \quad \Rightarrow \quad L_1(\lambda) = \lambda \begin{bmatrix} A & 0 \\ 0 & -C \end{bmatrix} + \begin{bmatrix} B & C \\ C & 0 \end{bmatrix},$$

$$v = e_m \quad \Rightarrow \quad L_2(\lambda) = \lambda \begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & C \end{bmatrix}.$$

Optimal κ_L Versus κ_P

Theorem 2 *Let λ be a simple eigenvalue of P. Then*

$$\frac{\inf_v \kappa_L(\lambda; v)}{\kappa_P(\lambda)} \le m^2 \rho.$$

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. 1:
  one of pencils with $v = e_1$ and $v = e_m$ is about as well conditioned as P itself for λ.

Quadratic Case: Scaling

$Q(\lambda) = \lambda^2 A + \lambda B + C$, $a = \|A\|_2, \quad b = \|B\|_2, \quad c = \|C\|_2$.

For optimality of $\kappa_L(\lambda)$ 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

- if problem is not too heavily damped,
- for elliptic QEPs.

Companion Linearizations

$$\textbf{L1}: \quad \lambda \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} B & C \\ -I & 0 \end{bmatrix}$$ **first companion**

$$\textbf{L2}: \quad \lambda \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} B & -I \\ C & 0 \end{bmatrix}$$ **second companion**

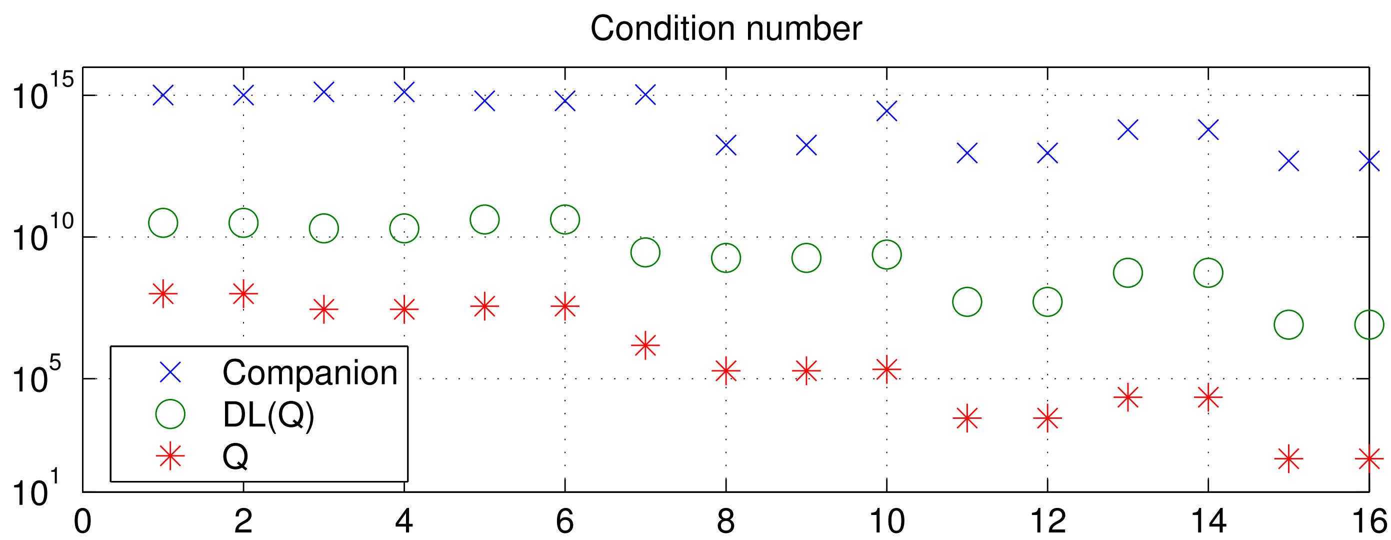
Further analysis shows these linearizations are

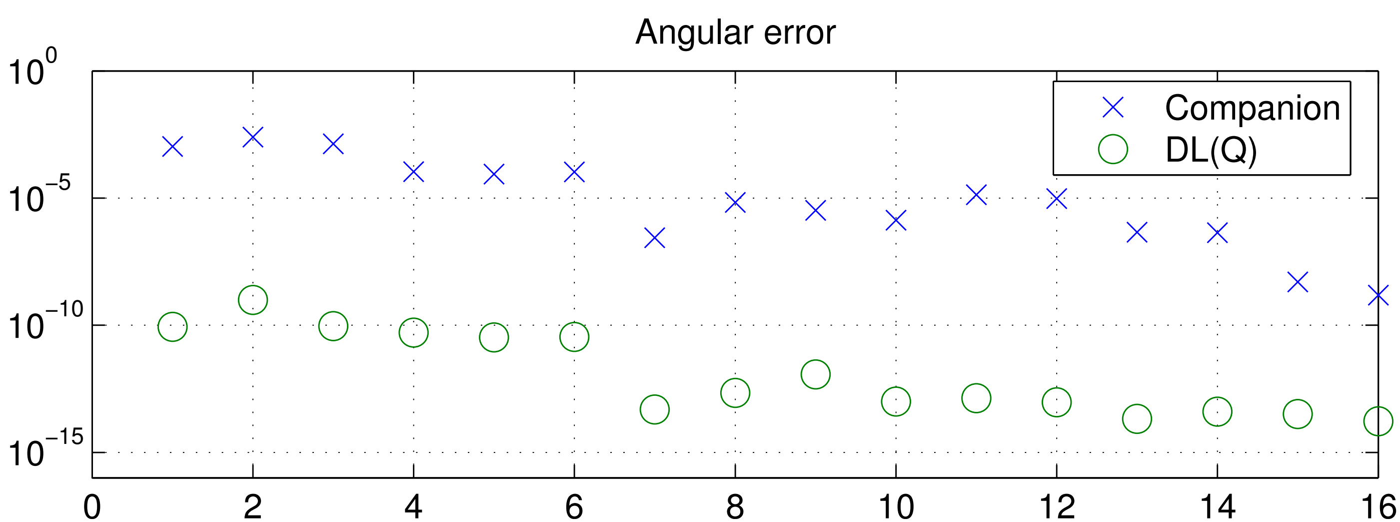
- as well conditioned as Q if $a \approx b \approx c \approx 1$,
- potentially **worse conditioned** than Q if left e'vec of L is much larger than left e'vec of Q.

Example 1: Nuclear Power Plant (1)

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

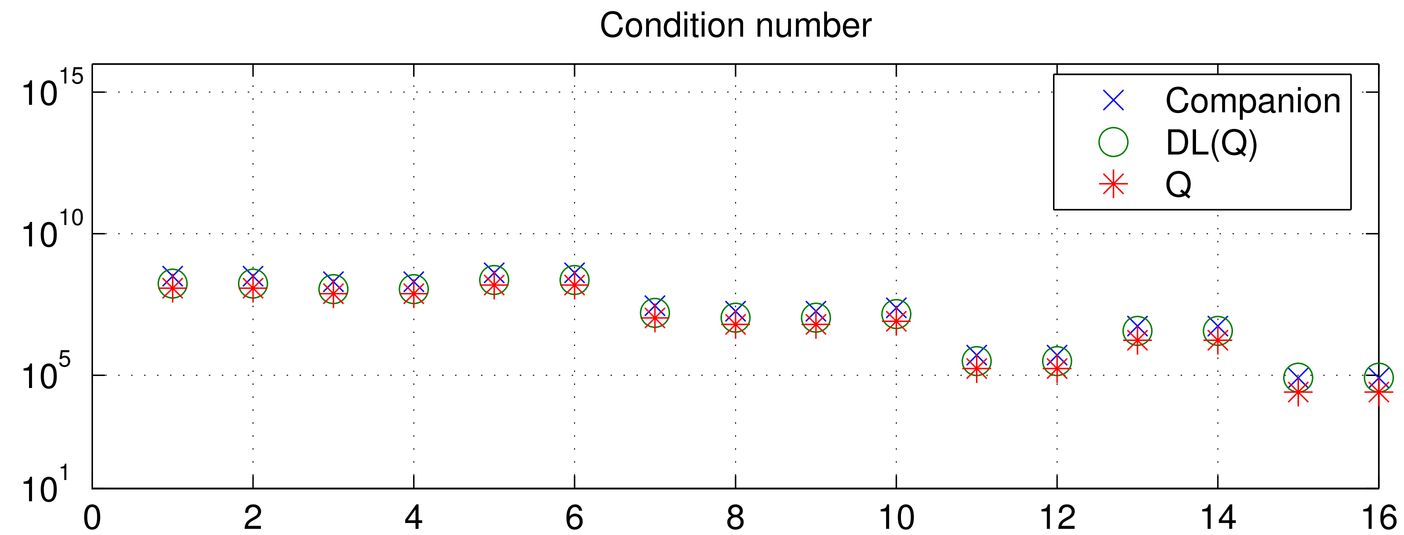
$|\lambda| \in (17, 362)$; Unscaled; $v = e_1$, $\rho = 7 \times 10^4$.

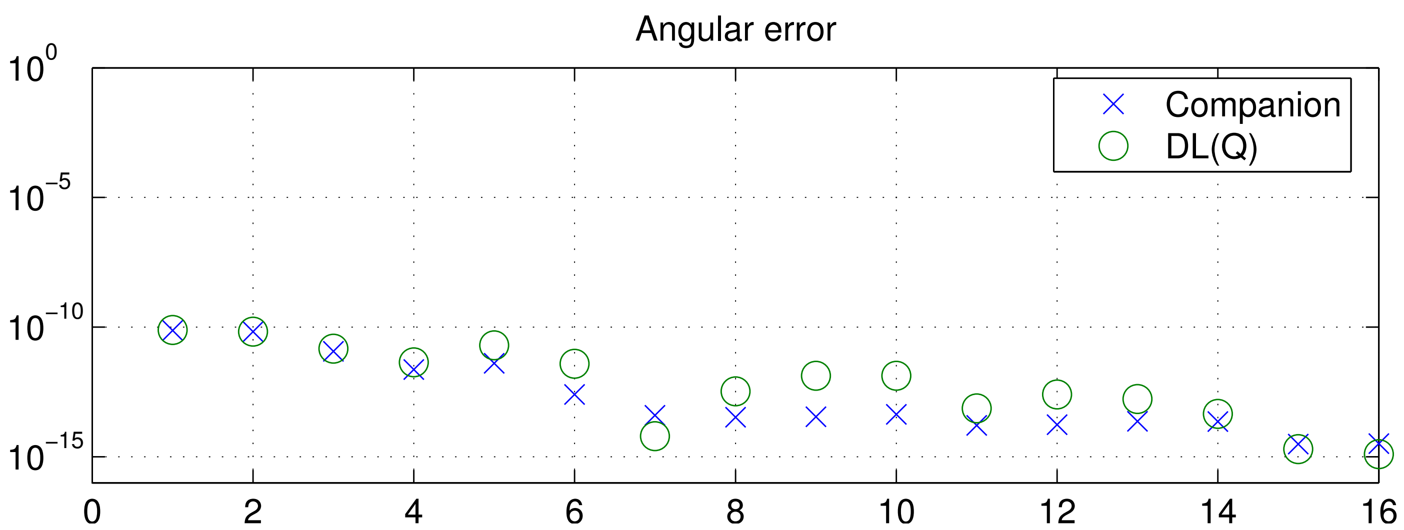




Example 1: Nuclear Power Plant (2)

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

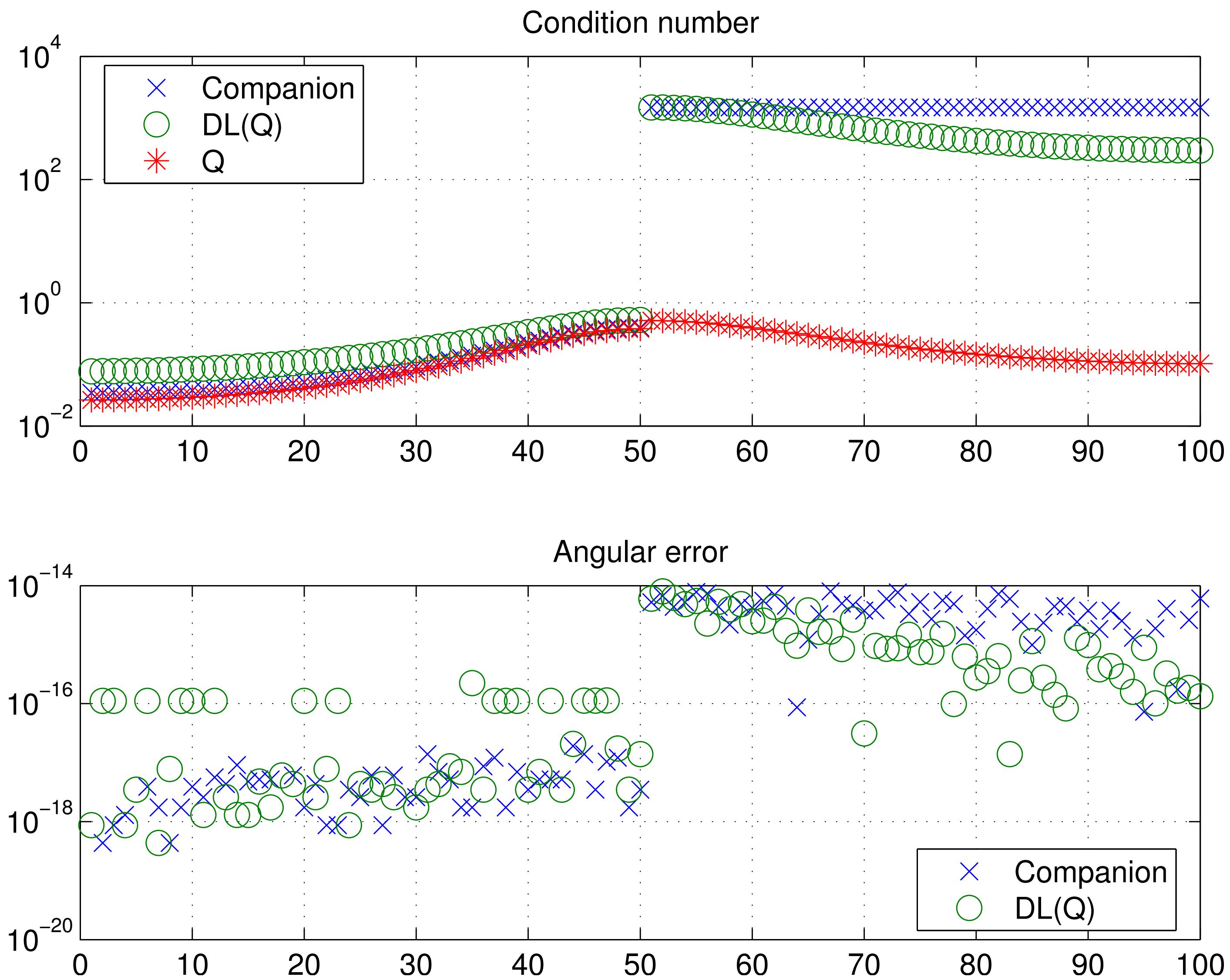




Example 2: Damped Mass-Spring (1)

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

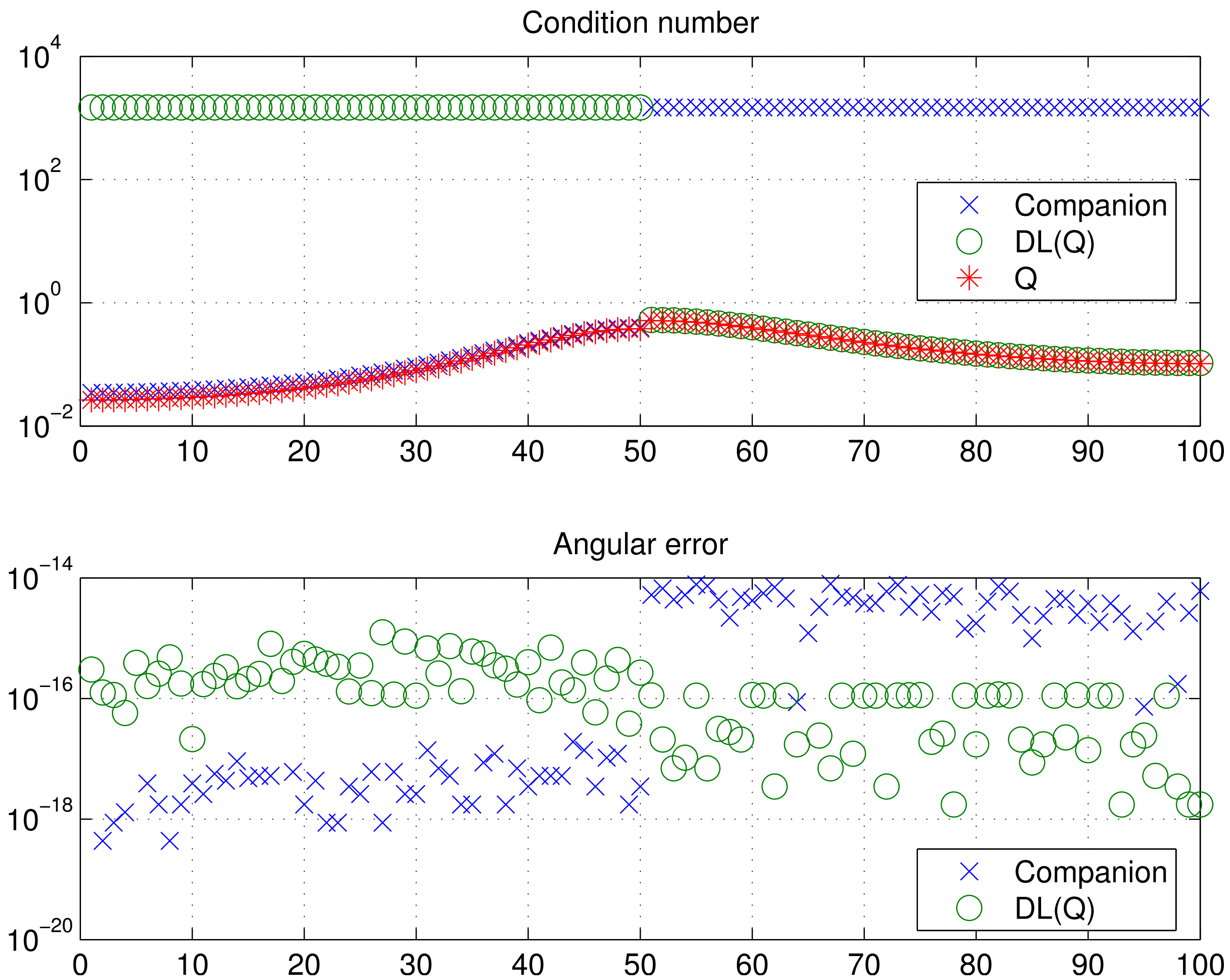
Unscaled; $v = e_1$, $\rho = 320$.



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

Solve ei'problem for P via linearized L:

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

- If $\max_i \|A_i\|_2 \approx \min(\|A_0\|_2, \|A_m\|_2)$ then L for $v = e_1$ and $v = e_m$ **optimally conditioned** within $\mathbb{DL}(P)$ 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.

Conclusions

Solve ei'problem for P via linearized L:

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

- If $\max_i \|A_i\|_2 \approx \min(\|A_0\|_2, \|A_m\|_2)$ then L for $v = e_1$ and $v = e_m$ **optimally conditioned** within $\mathbb{DL}(P)$ 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.

★ Justifies solving $P(\lambda)x = 0$ by linearization.

★ Guides choice of linearization.