Department of Computer Science
University of Illinois at Urbana-Champaign
http://www.cs.uiuc.edu

Center for Simulation of Advanced Rockets
University of Illinois at Urbana-Champaign

Improving the Convergence of the Jacobi-Davidson Algorithm

Eric de Sturler
✉ sturler@uiuc.edu http://www.cs.uiuc.edu/contacts/faculty/desturler.html

Seventh SIAM Conference on Applied Linear Algebra

North Carolina State University, Raleigh, NC, October 23-25, 2000

Overview

- The Jacobi-Davidson Algorithm
- Two Potential Problems
- Spurious Eigenvalues
- Poor Extension of Search Space
- Results
- Conclusions

Published in

- Iterative Methods in Scientific Computation II, 1999, IMACS. Proceedings IMACS Iterative Methods Conference, Austin 1998,
- Improving the Convergence of the Jacobi-Davidson Algorithm, Tech. Report UIUCDCS-R-2000-2173, June 2000 (submitted).

The Jacobi-Davidson Algorithm

Linear eigenvalue problem: $Ax = \lambda x$

Orthonormal basis for search space: V_k; $W_k = AV_k$; $H_k = V_k^* W_k \to (\vartheta, s)$

Ritz pair: (ϑ, u) where $u = V_k s$ and residual $r_k = Au - \vartheta u$

Solve: $(I - uu^*)(A - \vartheta I)(I - uu^*)t = -r, \qquad t \perp u$

$t = (I - V_k V_k^*)t / \|(I - V_k V_k^*)t\|$

$V_{k+1} = [V_k \,|\, t]; \qquad W_{k+1} = [W_k \,|\, At]; \qquad H_{k+1} = V_{k+1}^* W_{k+1};$

Continue

Two Potential Problems

Problem I.

Let $[s \,|\, S_c]$ be unitary matrix, $A_\vartheta = A - \vartheta I$,

Let $S_c^* V_k^* A_\vartheta V_k S_c = S_c^* H_k S_c - \vartheta I$ be ill-conditioned.

So (rel) small perturbation yields H_k with double eigenvalue ϑ, and corresponding eigenvector s cannot be computed to high accuracy.

This arises in cases where no other eigenvalues of A close to ϑ!

Second eigenvalue (close to) ϑ is spurious eigenvalue. Stagnation.

Problem II.

We solve correction equation to extend our search space.

Solve: $(I - uu^*)(A - \vartheta I)(I - uu^*)t = -r, \qquad t \perp u$

But solution is very close to search space: $\|(I - V_k V_k^*)t\|_2 \ll \|t\|_2$

The resulting vector $(I - V_k V_k^*)t$ contains mainly noise and will not provide an effective extension to the search space. Slow convergence.

Spurious Eigenvalues

Let $[s \,|\, S_c]$ be unitary and $S_c^* H_k S_c - \vartheta I = \Phi \Omega \Psi^*$ and ω_1/ω_{k-1} large.

$\|(S_c^* H S_c - \vartheta I)\psi_{k-1}\|_2 = \omega_{k-1}$

So, small (rel.) perturbation of H_k has multiple eigenvalue ϑ.
Eigenvector s of H_k associated with ϑ is ill-conditioned and is typically not computed to high accurracy: The method stalls.
The problem actually gets worse as the search space grows.

Main cause in JD is A_ϑ is ill-conditioned over range$(V_k S_c)$.
Either, because A is ill-conditioned, which doesn't mean x is, or approx. eigenvalue is more accurate than approx. eigenvector or v.v.

Another possibility, $A_\vartheta V_k S_c$ is not ill-conditioned, but for some y we have $A_\vartheta V_k S_c y \,\underset{\sim}{\perp}\, V_k S_c$. When this happens, the method stalls, but the problem may go away as the search space grows.

Spurious Eigenvalues

Solutions:

- Refined eigenvectors (Jia'95): $s = \arg\min_y \|(A - \vartheta I)V_k y\|_2$
  For Arnoldi (IRA) reduces to SVD for Hessenberg matrix, but ...
  This is not generally possible here! Expensive to do for $n \times k$ system.
  For JD, vectors in V_k do not span a Krylov space, so in general
  $\mathrm{range}((A - \vartheta I)V_k) \not\subseteq \mathrm{range}([V_k\ \tilde{v}])$. No way to make SVD cheap.

- Consider $S_c^* V_k^* (A - \vartheta I) V_k S_c = S_c^* H S_c - \vartheta I = \Phi\Omega\Psi$
  If $\omega_1/\omega_{k-1} > tol_1$, select largest $\omega_{k_1} \cdots \omega_{k_2}$ such that $\omega_{k_1}/\omega_{k_2} < tol_2$.
  Then $\hat{S} = [s \,|\, S_c\psi_{k_1} \cdots S_c\psi_{k_2}]$; $V = V_k\hat{S}$; $W = W_k\hat{S}$; $H = \hat{S}^* H_k \hat{S}$;
  Cheap implementation using Givens rotations ($O(Nk)$).
  Needed only few times in entire solve, whereas refined eigenvector requires $n \times k$ SVD at every step.

Spurious Eigenvalues

From paper Jia&Stewart (applied to our case) we know:
Ritz value ϑ converges unconditionally,
Ritz vector $u = V_k s$ converges if $\exists \varepsilon : \omega_{k-1} \equiv \sigma_{\min}(S_c^* H_k S_c - \vartheta I) \geq \varepsilon > 0$.

So adapt sequence of subspaces such that we maintain same converging Ritz value and that we keep the corresponding Ritz vector sufficiently well-conditioned.

This gives criteria when and how to truncate for chosen tolerance.

For computed $\tilde{s} : \|\tilde{s} - s\|_2 \lesssim \frac{\|H_k\|}{\omega_{k-1}} \varepsilon_{mach}$, hence we want $\frac{\|H_k\|_2}{\omega_{k-1}} < \frac{tol}{\gamma \varepsilon_{mach}}$
where *tol* is required accuracy (residual) and γ provides a *margin*.
If ϑ not largest eigenvalue or $\vartheta \ll \omega_1$ we use $\frac{\|H_k\|_2}{\omega_{k-1}} \approx \frac{\omega_1}{\omega_{k-1}} < \frac{tol}{\gamma \varepsilon_{mach}}$.

Also consider $\|r_s\|_2$ to judge effect of ill-conditioning and $\|r_s\|_2 \leq \|r\|_2$.

Poor Extension of Search Space

Solutions:

I.

Compute $A_{u,\vartheta} V_k S_c = QR$; now compute $C_m \perp Q$

$c_1 = (-r + QQ^* r)/\rho;\ \ \rho = \|-r + QQ^* r\|$

$A_{u,\vartheta} C_m = QQ^* A_{u,\vartheta} C_m + C_{m+1} \underline{G}_m$

Solve $t = V_k S_c y_1 + C_m y_2$ such that $\|-r - A_{u,\vartheta} t\|$ minimal

II.

We want to generate Krylov space orthogonal to V_k:

$c_1 = -r/\|r\|$ and $A C_m = V_k V_k^* A C_m + C_{m+1} \underline{G}_m$

Let $t = V_k S_c y_1 + C_m y_2$

Solve $-r - (I - uu^*) A_\vartheta (I - uu^*) t \perp A_{u,\vartheta} [V_k S_c \,|\, C_m]$

Expensive (orthogonalization), but only at few iterations.

Results

Problem 1:

$A = SDS^{-1}$, where $D = diag((k/100)^2 - 0.8)$, $k = 1\ldots100$, and S is bidiagonal with 0.8 on diagonal and 1 super diagonal. smallest absolute $\lambda = -0.0079$ (interior). 10 steps of GMRES for correction equation. Problem from Sleijpen&v/dVorst with scaling to make problem non-symmetric.

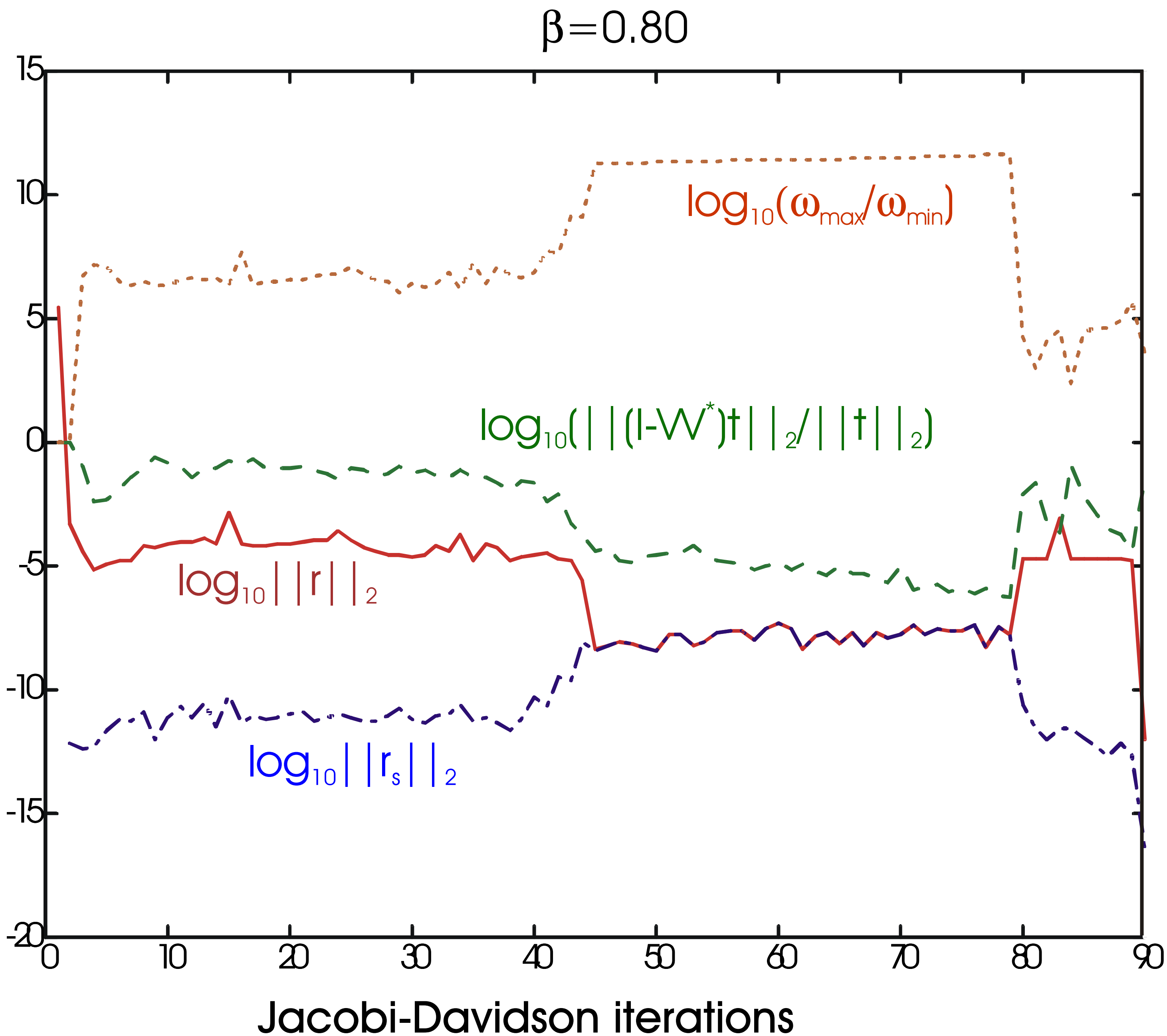
Problem 2:

$-u_{xx} - u_{yy} + 20u_x - 30u_y = 0$ on unit square with Dirichlet b.c.
$x = 0, y = 1 : u = 1$ and $x = 1, y = 0 : u = 0$. 22x22 grid.
20 steps of GMRES for correction equation.

Problem 3:

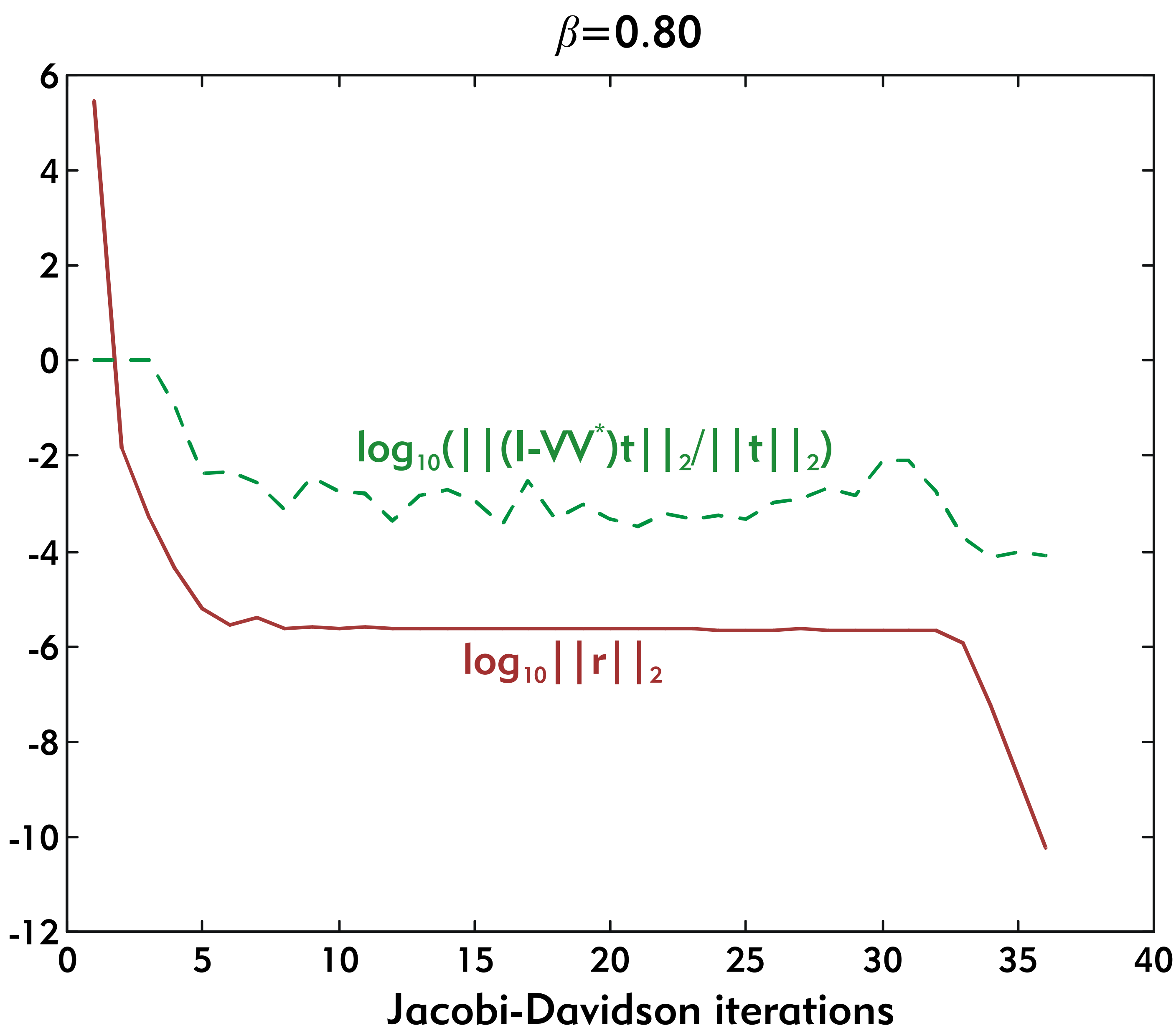
West0479; solve for $\lambda \approx (-17.825, -4.6376)$. 20 steps of GMRES for correction equation.

Results (problem 1)



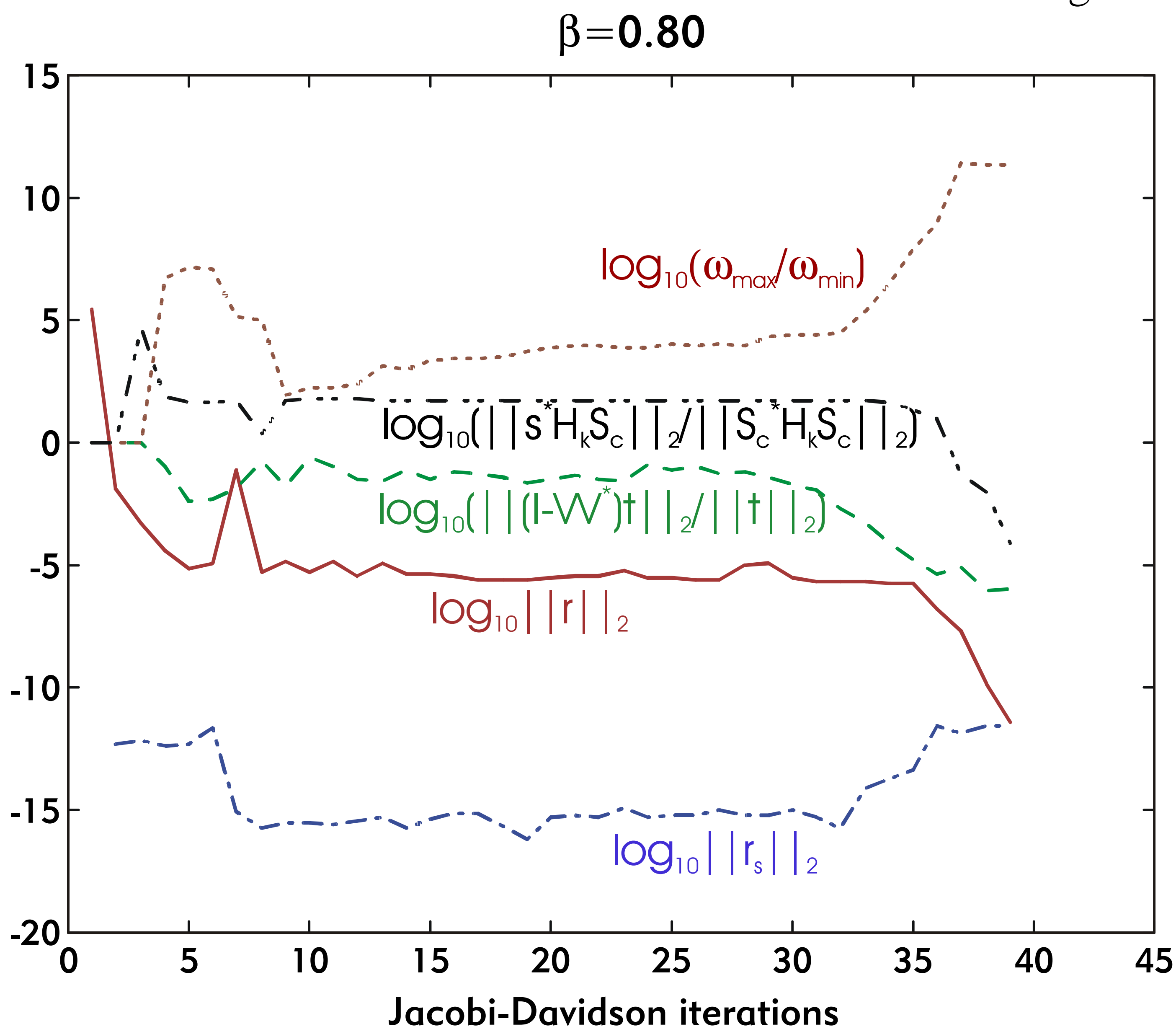
Results (problem 1)

Using refined eigenvectors (expensive)



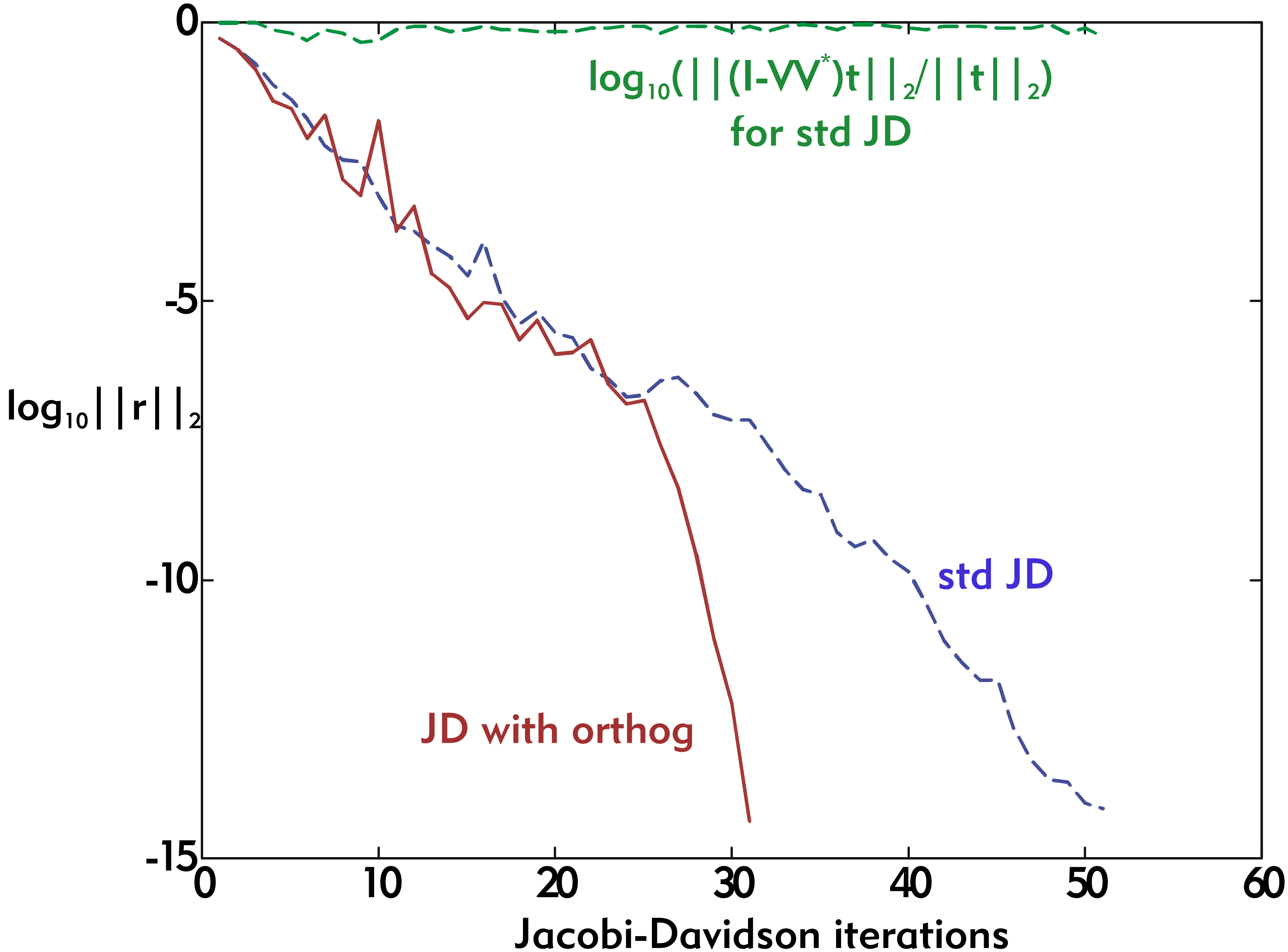
Results (problem 1)

Jacobi-Davidson with truncation for well-cond. eigenvector

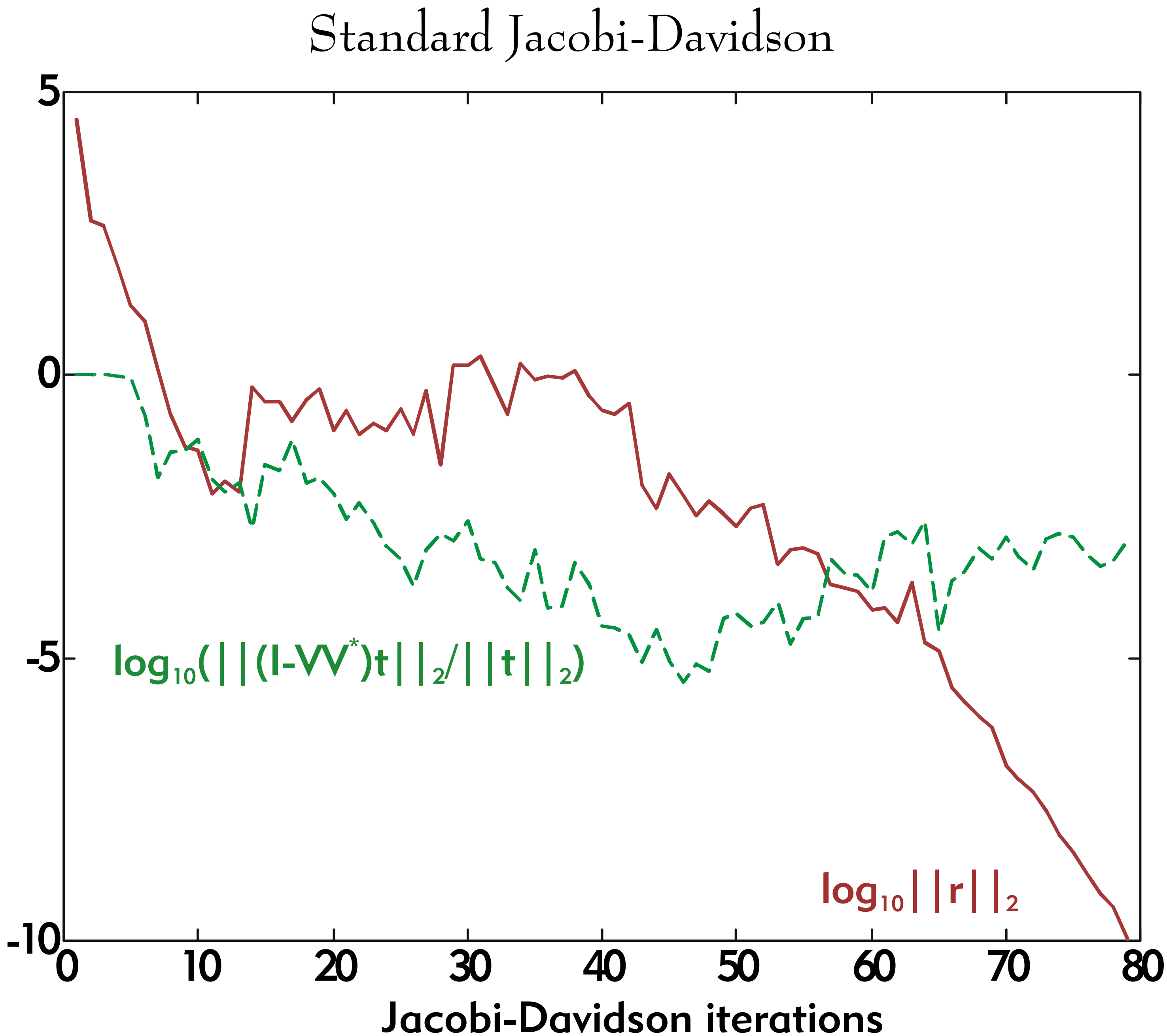


Results (problem 2)

JD with orthogonalization uses method 1 at every step

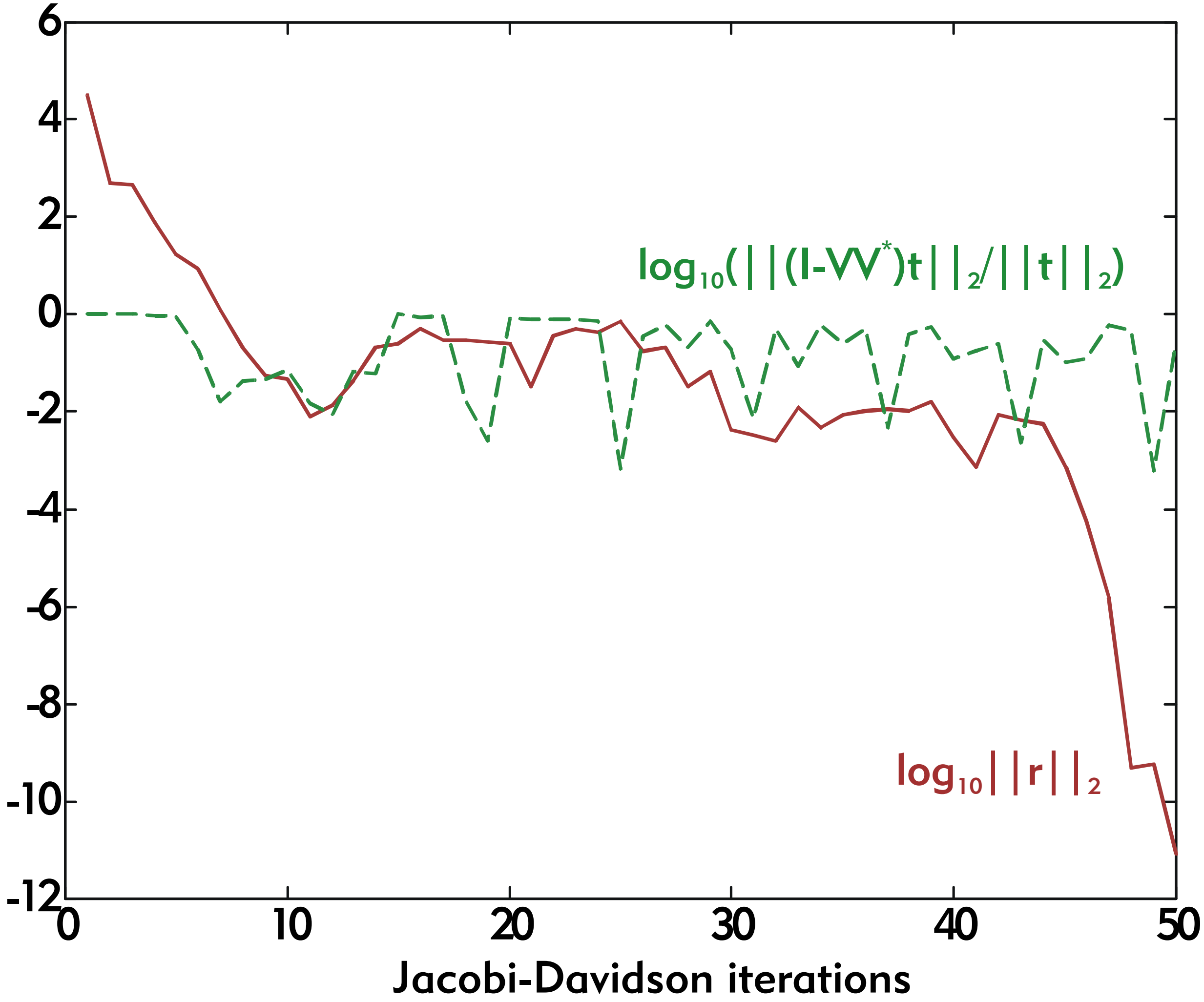


Results (problem 3)

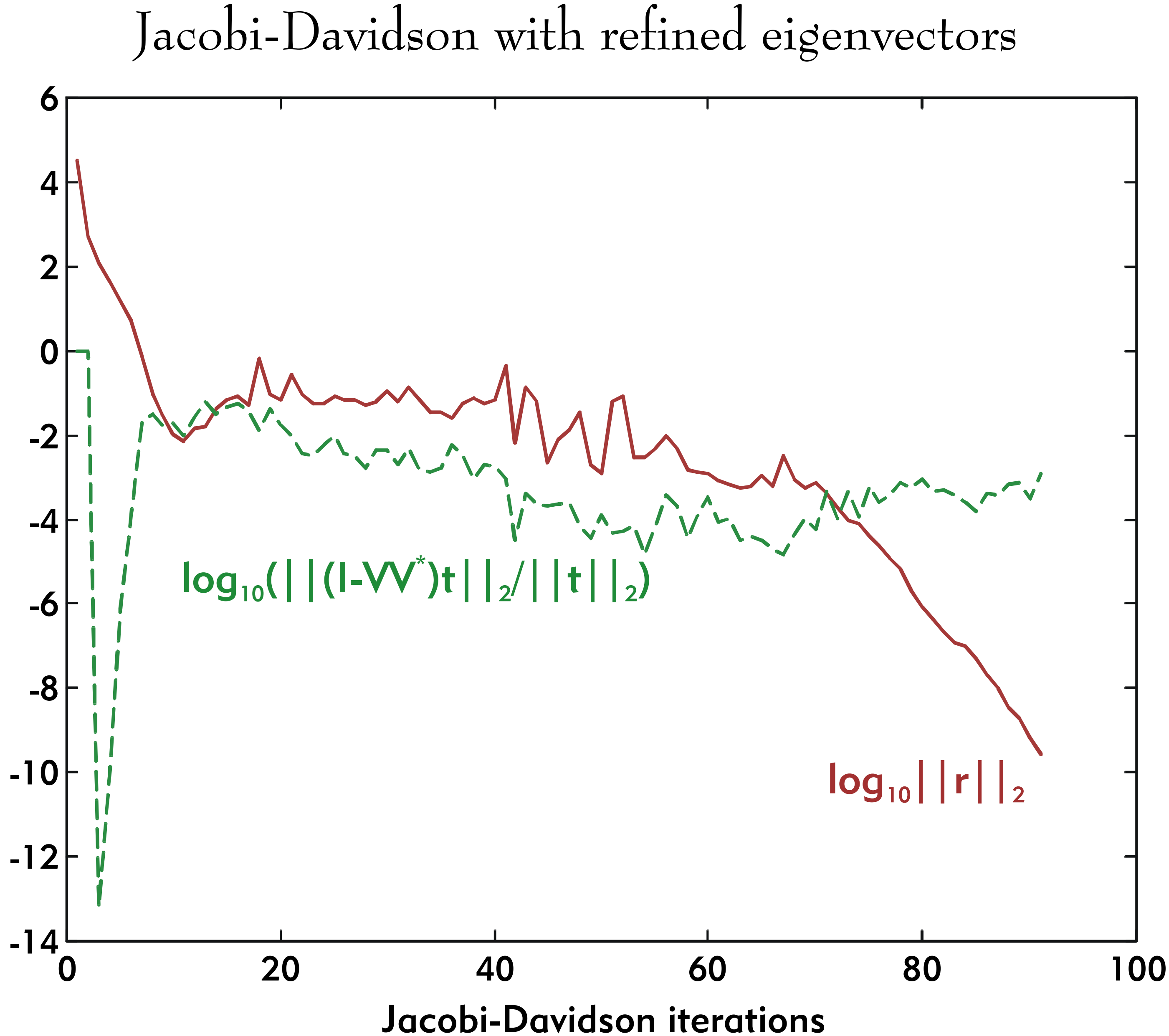


Results (problem 3)

JD with orthogonalization (method 1) only when necessary



Results (problem 3)



Conclusions

- The JD algorithm may suffer from slow convergence (2 problems):
  - Spurious eigenvalues due to ill-conditioning shifted operator over search space
  - Extension/solution to correction equation close to search space
- Truncating to improve conditioning of Ritz vector works very well; cheaper than refined Ritz vector in context of JD method.
- Preliminary tests: faster convergence than refined eigenvectors and smaller residual.
- Techniques to improve independence of correction equation solution leads to faster convergence
- Both strategies should be applied only iteration-wise when needed (for efficiency).
- Combine strategies with refined eigenvectors?