A Grassmann-Rayleigh Quotient Iteration for Computing Invariant Subspaces

2002 SIAM Student Paper Prize presentation

Pierre-Antoine Absil (1)
Rodolphe Sepulchre (1)
Robert Mahony (2)
Paul Van Dooren (3)

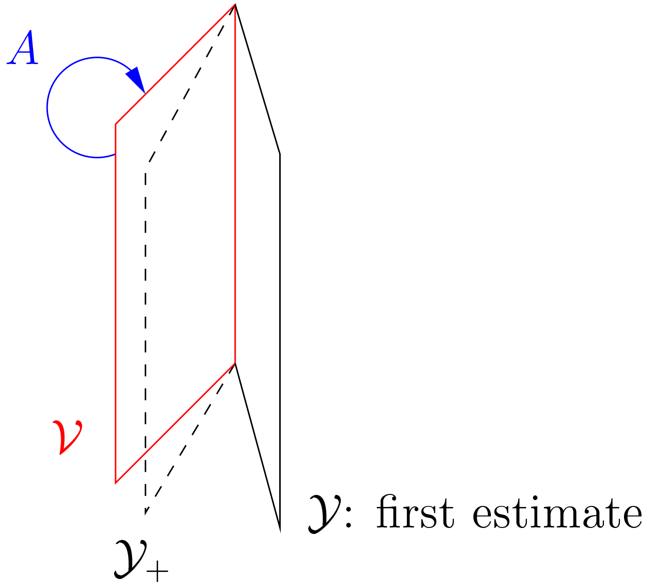
(1) Department of Electrical Engineering and Computer Science,
Université de Liège, 4000 Liège Sart-Tilman, Belgium.
{absil,sepulchre}@montefiore.ulg.ac.be
(2) Department of Engineering, Australia National University, ACT, 0200, Australia.
(3) Department of Mathematical Engineering, Université Catholique de Louvain, Belgium.

Computing invariant subspaces

Let A be n-by-n matrix, $A = A^T$.

Objective: find an eigenspace (invariant subspace) of dimension p.

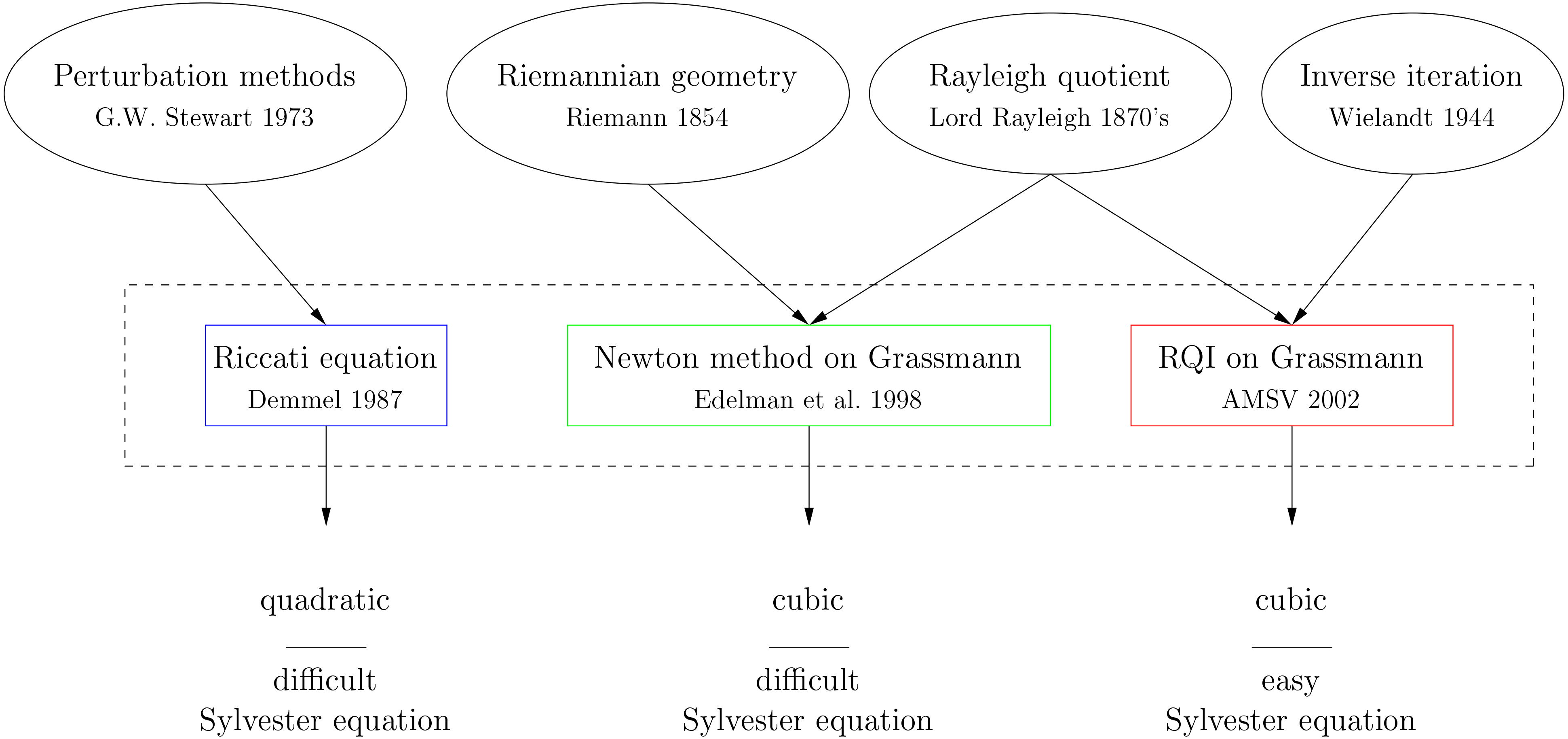
$\mathcal{V}$: $A\mathcal{V} \subset \mathcal{V}$.



Applications: numerous! (pattern recognition, antenna array processing, model reduction). Usually, $p << n$.

Outline

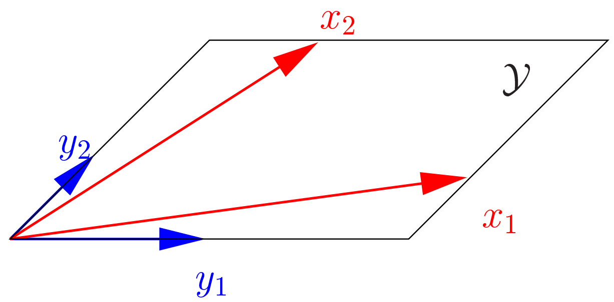
Algorithms that refine eigenspace estimates.



Representing subspaces

Subspace represented by basis.

Infinitely many possible choices.



Let $Y := [y_1|\dots|y_p]$ and $X := [x_1|\dots|x_p]$.

Then $\mathcal{Y} = \operatorname{span}(Y) = \operatorname{span}(X)$.

Equivalence classes of bases:

$\operatorname{span}(X) = \operatorname{span}(Y)$ iff $X = YM$ with M p-by-p invertible.

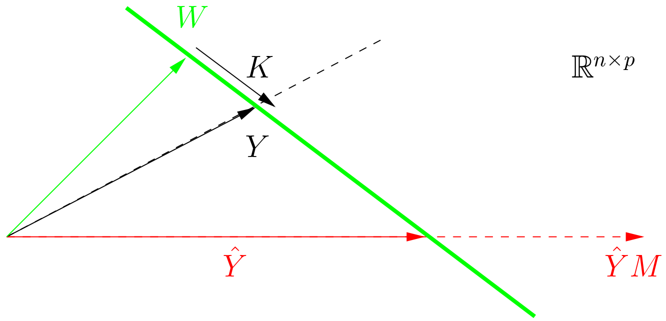
Grassmann manifold $\operatorname{Grass}(p, n)$: the set of the p-dimensional subspaces of $\mathbb{R}^n$.

Riccati equation

Y spans an eigenspace iff $\Pi_{Y_\perp} AY = 0$ where $\Pi_{Y_\perp} := (I - Y(Y^TY)^{-1}Y^T)$.

Problem: $\hat{Y}$ solution $\Rightarrow$ $\hat{Y}M$ solution for all M invertible.

Remedy: work in an affine section.



Then $\Pi_{Y_\perp} AY = 0$ becomes

$$\underbrace{W_\perp^T A W_\perp}_{A_{22}} K - K \underbrace{W^T A W}_{A_{11}} + \underbrace{W_\perp^T A W}_{A_{21}} - K \underbrace{W^T A W_\perp}_{A_{12}} K = 0.$$

Newton iteration: $\mathrm{span}(W + W_\perp K) \mapsto \mathrm{span}(W + W_\perp K_+)$

$$(A_{22} - KA_{12})\, K_+ \; - \; K_+\, (A_{22} - KA_{12}) = -A_{21} - KA_{12}K.$$

Newton method on Grassmann

Cost function = Rayleigh quotient

$$\rho(\mathrm{span}(Y)) := \mathrm{trace}((Y^TY)^{-1}Y^TAY).$$

Riemannian metric on Grassmann.

Gradient: $(\mathrm{grad}\rho)_{\diamond Y} = \Pi_{Y_\perp} AY = 0$ iff Y spans eigenspace.

Hessian: $(\nabla_N \mathrm{grad}\rho)_{\diamond Y} = \Pi_{Y_\perp} AN_{\diamond Y} - N_{\diamond Y}Y^TAY$.

Newton equation: $(\nabla_N \mathrm{grad}\rho)_{\diamond Y} = -(\mathrm{grad}\rho)_{\diamond Y}$, i.e.

$$\Pi_{Y_\perp} AN_{\diamond Y} - N_{\diamond Y}Y^TAY = -\Pi_{Y_\perp} AY, \; Y^TN_{\diamond Y} = 0.$$

Iteration: $\mathrm{span}(Y) \mapsto \mathrm{span}(Y + Y_\perp K)$

$$Y_\perp^T AY_\perp K - KY^TAY = -Y_\perp^T AY.$$

Convergence rate: cubic.

RQI (Rayleigh Quotient Iteration)

Rayleigh quotient:

$$\rho(y) = \frac{y^T A y}{y^T y}.$$

Quadratic approximation of eigenvalues: Let (λ, v) be an eigenpair of A. Then

$$\rho(v + \epsilon) = \lambda + O(\epsilon^2).$$

Rayleigh quotient iteration: $\mathrm{span}(y) \mapsto \mathrm{span}(z)$,

$$z = (A - \rho(y)I)^{-1} y.$$

Convergence rate: cubic convergence to the eigenvectors of A.

$\boxed{\text{Grassmann-RQI}}$

$$\overset{n}{A}\;\overset{1}{z} - \frac{y^T A y}{y^T y}\,\overset{1}{z} = \overset{1}{y}$$

$$\overset{1}{y} \rightsquigarrow \overset{p}{Y}$$

$$\overset{n}{A}\;\overset{p}{Z} - \overset{p}{Z}\;\overset{p}{(Y^T Y)^{-1}}\;\overset{p}{Y^T A Y} = \overset{p}{Y}$$

$\boxed{\text{Grassmann-RQI}}$

$$AZ - Z(Y^TY)^{-1}Y^TAY = Y$$

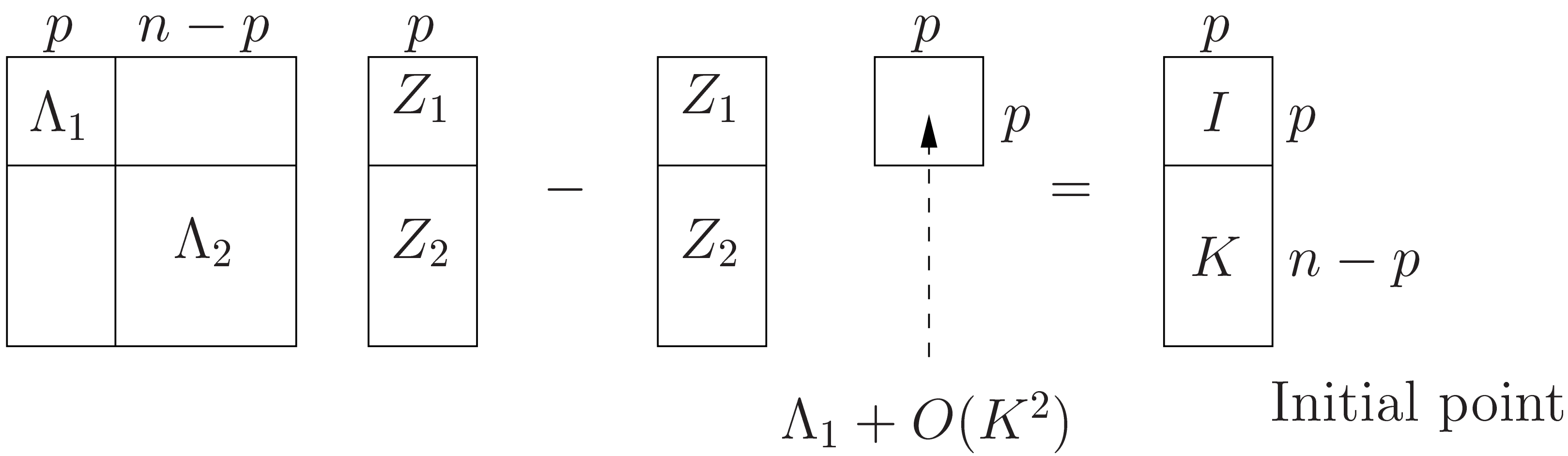
Grassmann-RQI: span$(Y) \mapsto$ span(Z).

$$AZM - ZM((YM)^T(YM))^{-1}(YM)^TAYM = YM$$

Recall: Y and Y' have the same span iff $Y' = YM$ with M invertible.

Grassmann-RQI : cubic

Take A diagonal. Target: span of $\begin{bmatrix} I \\ 0 \end{bmatrix}$.



Results: $Z_2 = O(K)$ (easy) and $Z_1^{-1} = O(K^2)$ (harder).

New subspace $=$ span $\left[\begin{smallmatrix} Z_1 \\ Z_2 \end{smallmatrix}\right]$ $=$ span $\left[\begin{smallmatrix} I \\ Z_2 Z_1^{-1} \end{smallmatrix}\right]$ $=$ span $\left[\begin{smallmatrix} I \\ O(K^3) \end{smallmatrix}\right]$.

From one-dimensional to p-dimensional iterates

Alg. on projective space $\mathbb{RP}^{n-1}$ Compute an eigenvector of A	Alg. on Grassmann manifold $\mathrm{Grass}(p,n)$ Compute a p-dim eigenspace of A
Power Method: $\boldsymbol{z} = A\boldsymbol{y}$, $\boldsymbol{y}_+ = \alpha \boldsymbol{z}$	Direct subspace iteration: $Z = AY$, $Y_+ = ZM$
(Shifted) Inverse Iteration: $\boldsymbol{z} = (A - \sigma I)^{-1}\boldsymbol{y}$, $\boldsymbol{y}_+ = \alpha \boldsymbol{z}$ i.e. $(A-\sigma I)\boldsymbol{z} = \boldsymbol{y}$. Scalar σ fixed.	Inverse subspace iteration: $Z = (A - \sigma I)^{-1} Y$, $Y_+ = ZM$. Scalar σ fixed.
Rayleigh quotient iteration: $(A - \rho_A(\boldsymbol{y})I)\boldsymbol{z} = \boldsymbol{y}$, $\boldsymbol{y}_+ = \alpha \boldsymbol{z}$, where $\rho_A(\boldsymbol{y}) = (\boldsymbol{y}^T A \boldsymbol{y})/(\boldsymbol{y}^T \boldsymbol{y})$.	Grassmann-RQI: $AZ - Z(Y^TY)^{-1}Y^TAY = Y$, $Y_+ = ZM$.

Practical implementation of GRQI (1)

Key equation:

$$AZ - Z(Y^TY)^{-1}Y^TAY = Y.$$

- Reduce A to a condensed form before starting the iteration (e.g. tridiagonal).
- Take the basis Y orthonormal:

$$AZ - ZY^TAY = Y.$$

$\boxed{\text{Practical implementation of GRQI (2)}}$

Key equation:

$$A\,Z \;-\; Z\,Y^T A Y \;=\; Y.$$

Diagonalize $Y^T AY$ (small, p-by-p) by rotating Y, $\to YQ$:

$$A\,ZQ \;-\; ZQ\;Q^T Y^T A Y Q \;=\; YQ.$$

Solve the p decoupled equations.

$\boxed{\text{Numerical cost}}$

Newton for $\Pi_{Y_\perp} AY = 0$ in affine section:
$\mathrm{span}(W + W_\perp K) \mapsto \mathrm{span}(W + W_\perp K_+)$

$$(A_{22} - KA_{12})\, K_+ \;-\; K_+\, (A_{22} - KA_{12}) = -A_{21} - KA_{12}K.$$

Coordinate-free Newton: $\mathrm{span}(Y) \mapsto \mathrm{span}(Y + Y_\perp K)$

$$Y_\perp^T AY_\perp \; K - K \; Y^T AY = -Y_\perp^T AY.$$

GRQI: $\mathrm{span}(Y) \mapsto \mathrm{span}(Z)$

$$A \; Z \;-\; Z \; Y^T AY \;=\; Y.$$

The big matrix $Y_\perp^T AY_\perp$ changes at each iteration step.

Newton vs GRQI - equations in Y coordinates

Work in coordinate system that makes $Y = \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right]$.

Decompose $A = \left[\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right]$.

Newton: $\operatorname{span}\left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right] \mapsto \operatorname{span}\left[\begin{smallmatrix} I \\ K \end{smallmatrix}\right]$

$$A_{22}\ K - K\ A_{11} = -A_{21}.$$

GRQI: $\operatorname{span}\left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right] \mapsto \operatorname{span}\left[\begin{smallmatrix} I \\ K \end{smallmatrix}\right]$

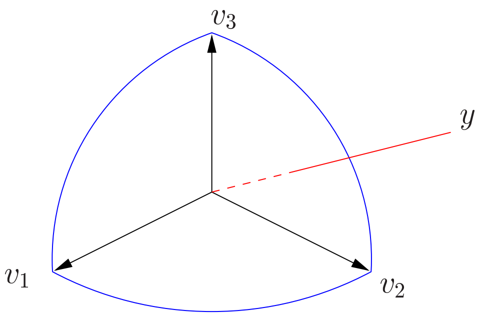
$$A_{11} + A_{12}K - Z_1 A_{11} Z_1^{-1} = Z_1^{-1}$$

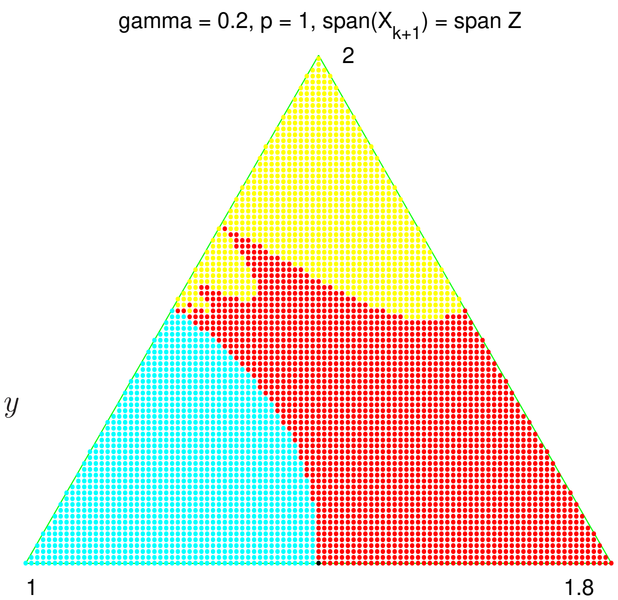
$$A_{22}\ K - K\ Z_1 A_{11} Z_1^{-1} = -A_{21}$$

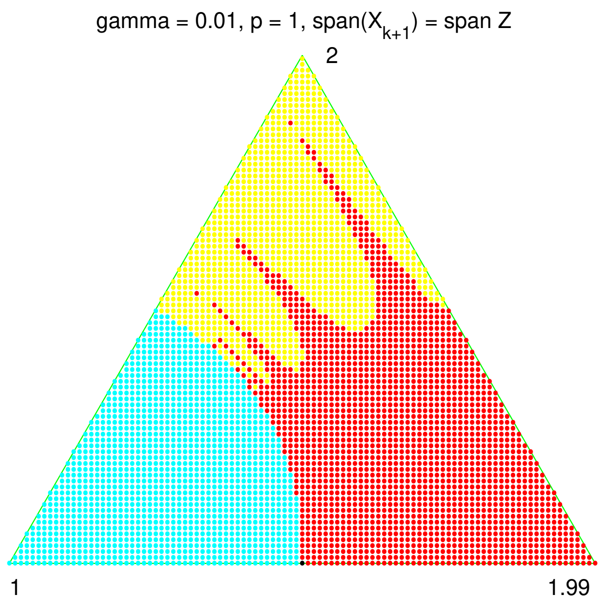
When $p = 1$, Newton and GRQI are identical.

Newton and GRQI - global behaviour ($p = 1$)

Case $p = 1$: Newton = GRQI.

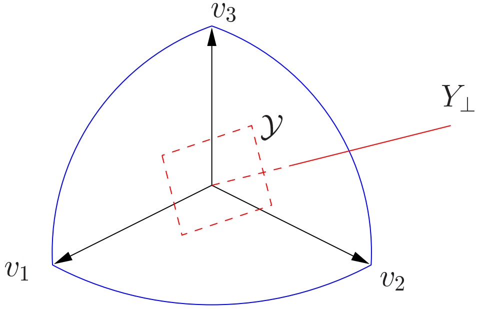


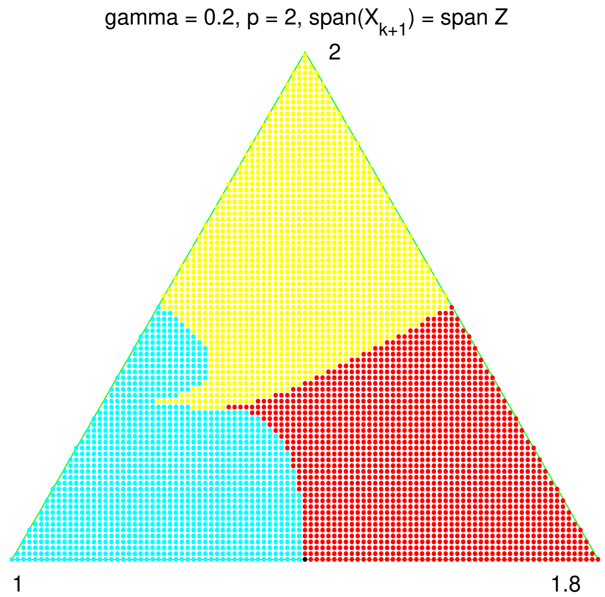


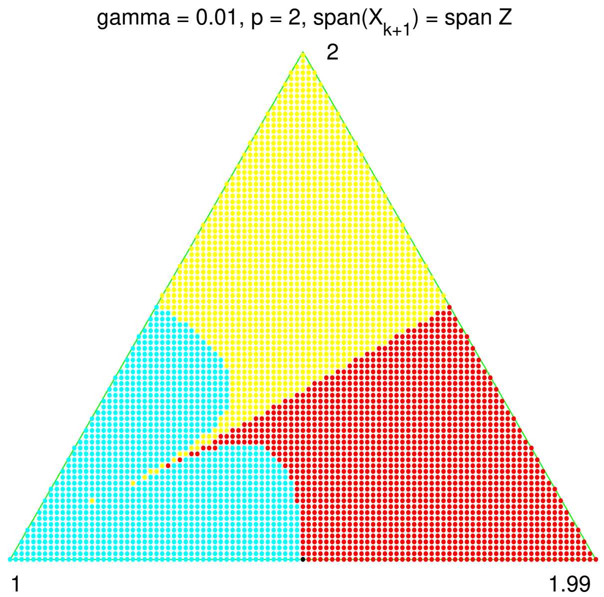


Newton and GRQI - global behaviour ($p > 1$)

Case $p > 1$: Newton $\neq$ GRQI. Basins of GRQI:



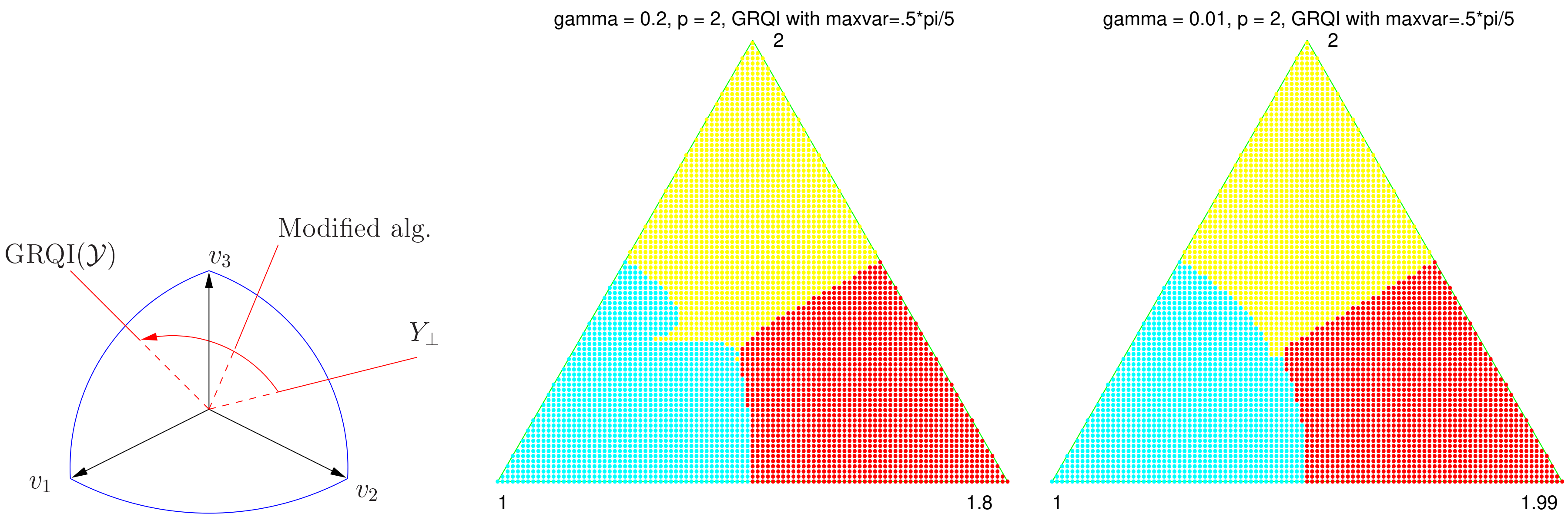




Basins of Newton: see previous slide (duality property!).

GRQI - improving global behaviour

Principle: put a threshold on the angles between two GRQI iterates.



Summary

Algorithms: GRQI (Grassmannian Rayleigh Quotient Iteration) and Newton method (on Grassmann for the Rayleigh Quotient cost function).

Task: refine estimates of eigenspaces of a symmetric matrix.

Convergence rate: cubic.

Numerical cost: GRQI takes advantage of structured A.

Global behaviour: different.

$\boxed{\text{Current work}}$

Globally convergent GRQI and Newton methods.

Extension of GRQI to nonsymmetric case, generalized eigenvalue problem, Hamiltonian pencils...