Invariant Subspace Computation - A Geometric Approach

Pierre-Antoine Absil

`PhD Defense`

`Université de Liège`

`Thursday, 13 February 2003`

Acknowledgements

Advisor:

Prof. R. Sepulchre
(Université de Liège)

Collaborations with:

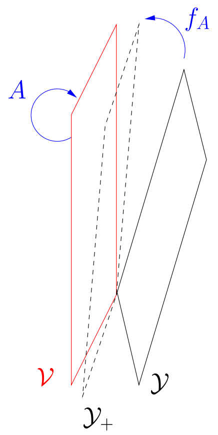
Prof. P. Van Dooren
(Université Catholique de Louvain)

Prof. R. Mahony
(Australian National University)

Outline

- Eigenspace computation: definitions.
- Eigenspace computation: an application.
- Geometric approach to eigenspace computation.
- New numerical algorithms:
  - Newton methods.
  - Shifted inverse iterations.
  - Global behaviour and improvements.

Topic: Refining estimates of eigenspaces



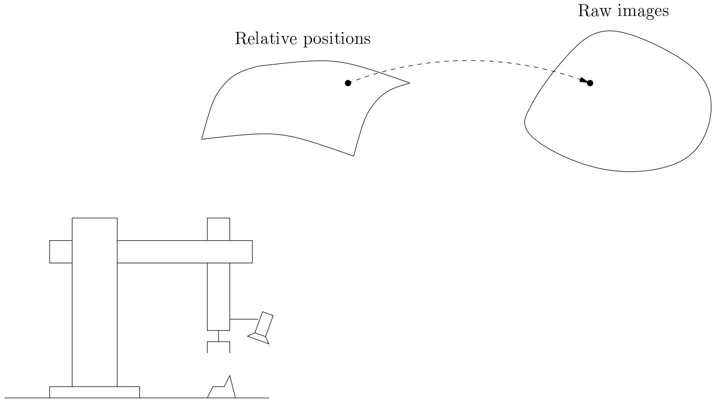
Data:

- n-by-n matrix A
- p-dimensional subspace $\mathcal{Y}$, approx. of eigenspace $\mathcal{V}$.

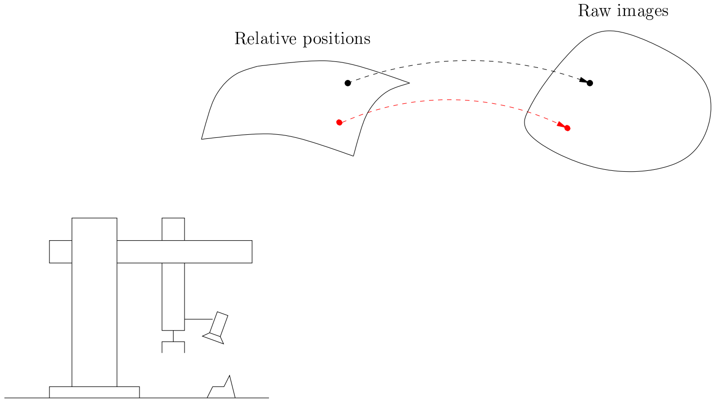
Task: compute better estimate of $\mathcal{V}$.

Definition of "eigenspace": $A\mathcal{V} \subseteq \mathcal{V}$.

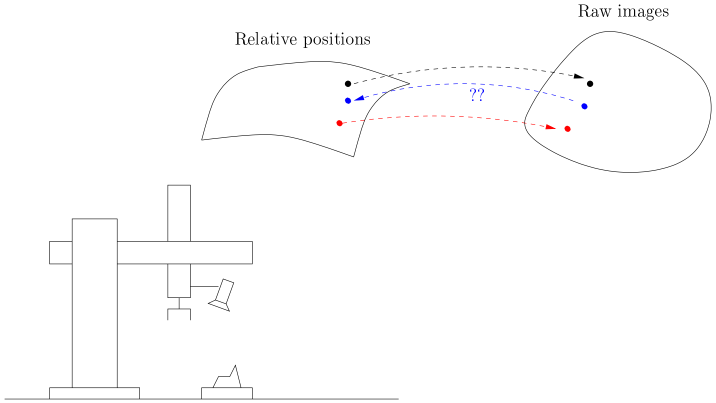
Application: visual positioning



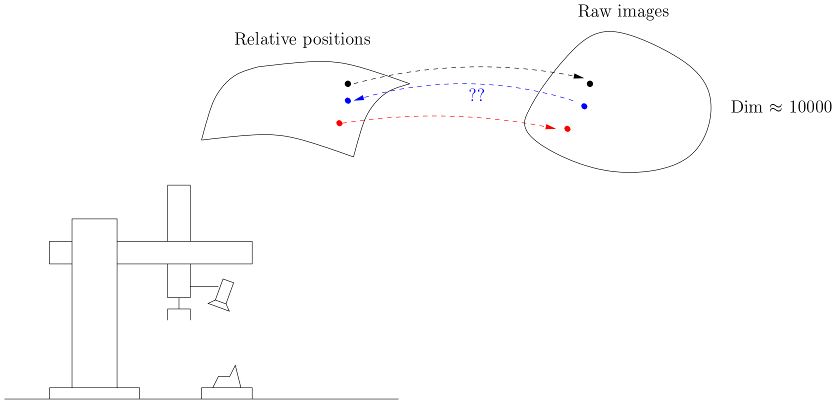
Application: visual positioning



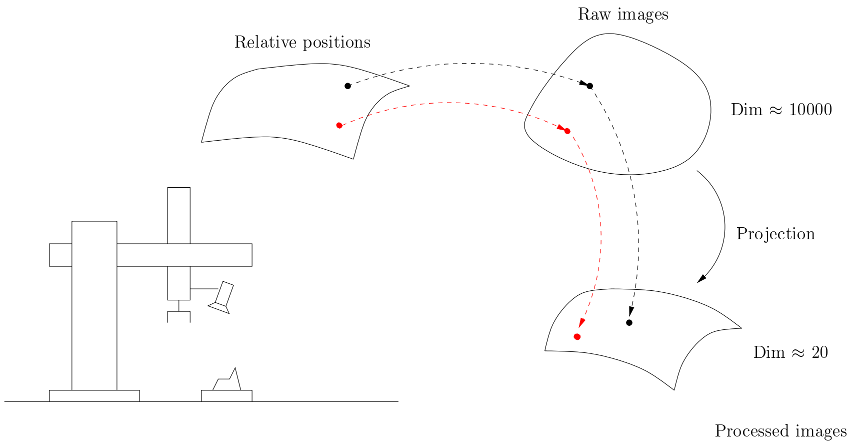
Application: visual positioning



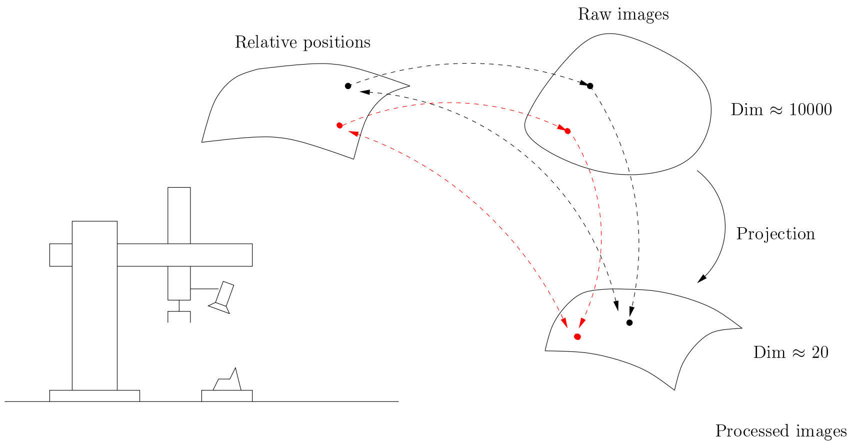
Application: visual positioning



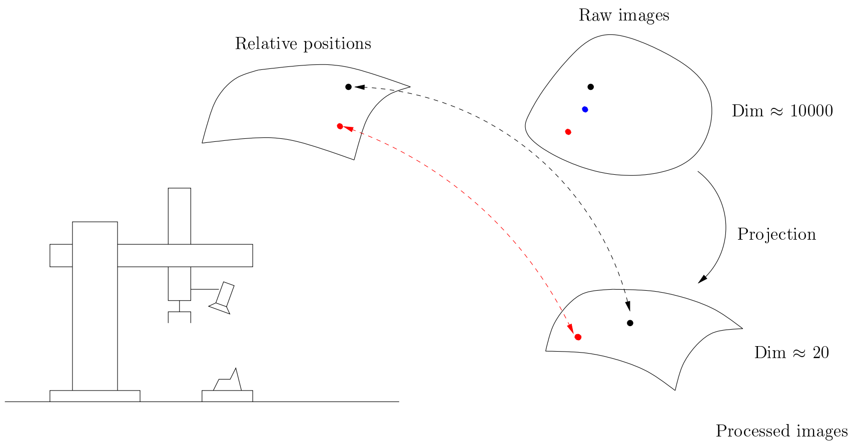
Application: visual positioning



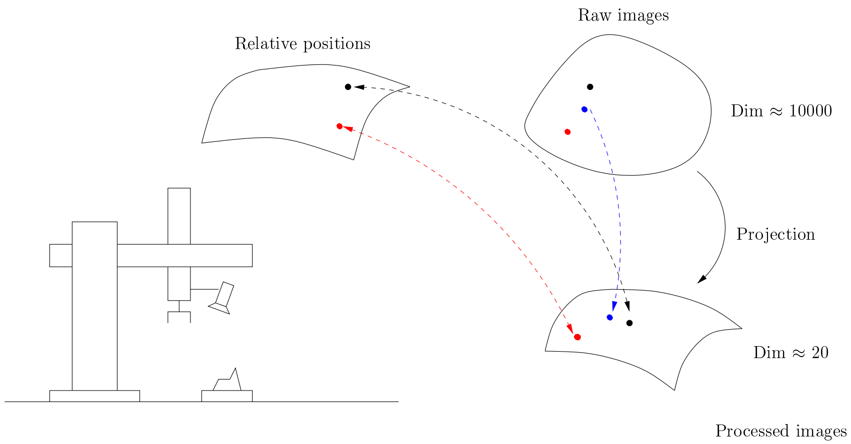
Application: visual positioning



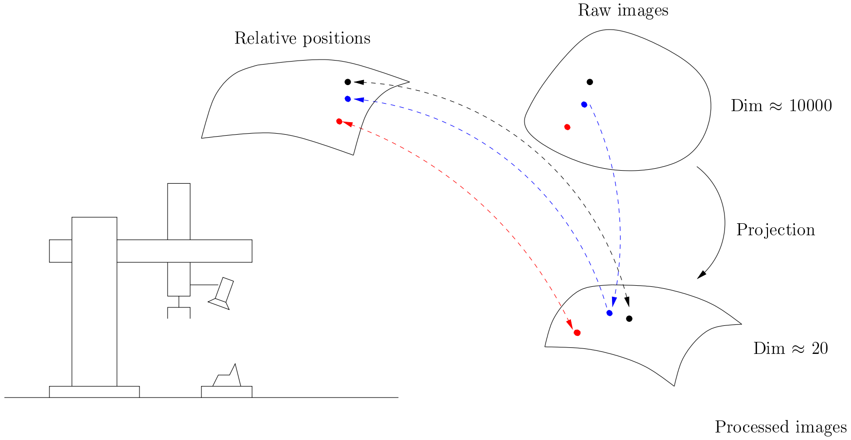
Application: visual positioning



Application: visual positioning

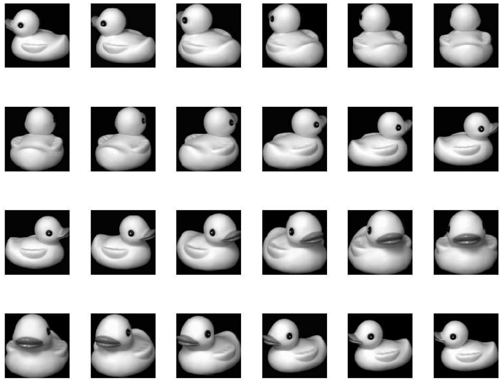


Application: visual positioning



Application: visual positioning

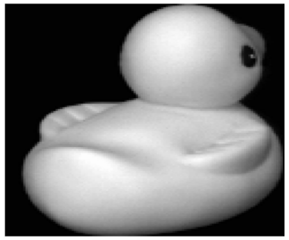
Learning...

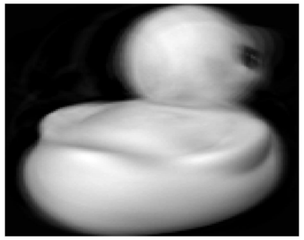


Compute covariance matrix and dominant eigenspace...

Application: visual positioning

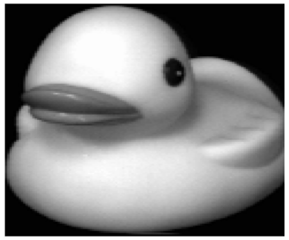
Raw image | Projection onto eigenspace.

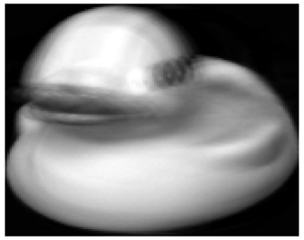




Application: visual positioning

Raw image | Projection onto eigenspace.

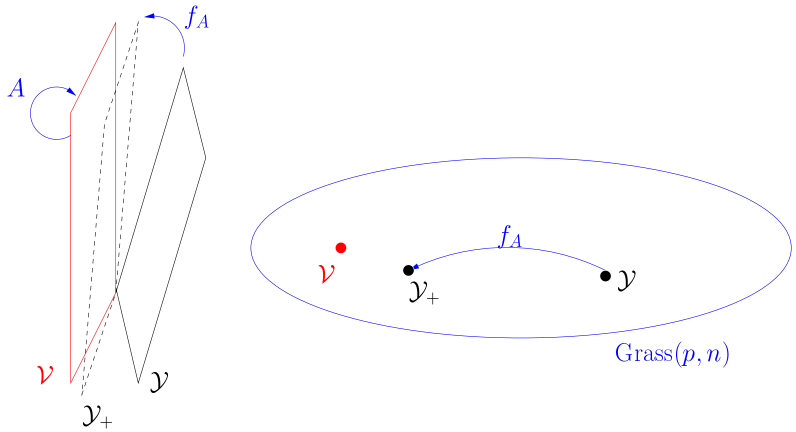




Outline

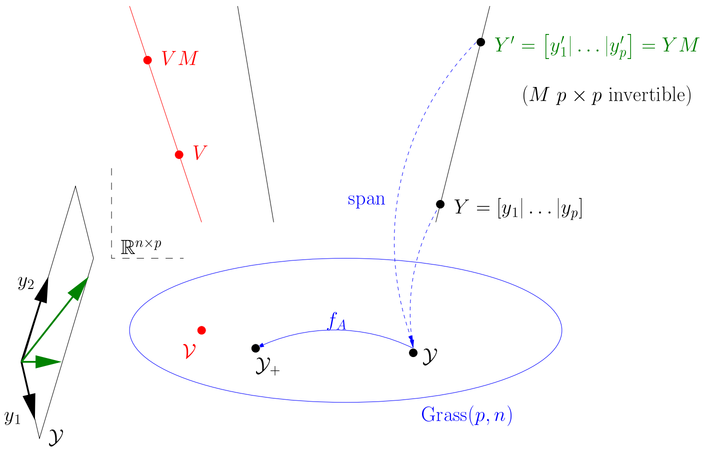
- Eigenspace computation: definitions.
- Eigenspace computation: an application.
- Geometric approach to eigenspace computation.
- New numerical algorithms:
  - Newton methods.
  - Shifted inverse iterations.
  - Global behaviour and improvements.

Refining estimates of eigenspaces

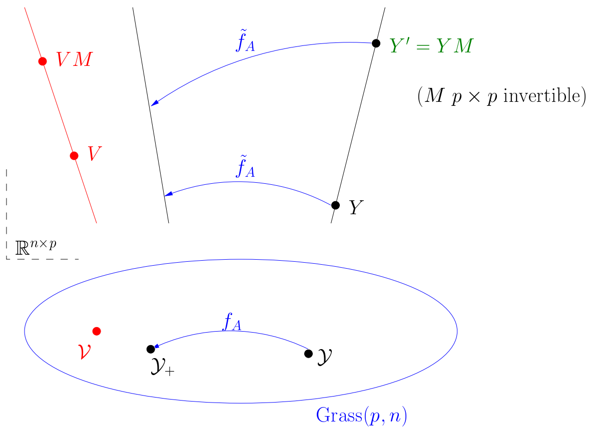


$p = \dim(\mathcal{V}) = \dim(\mathcal{Y}) = \dim(\mathcal{Y}_+)$.
Grass(p, n) (Grassmann manifold): the set of p-planes of $\mathbb{R}^n$.

Matrix representations



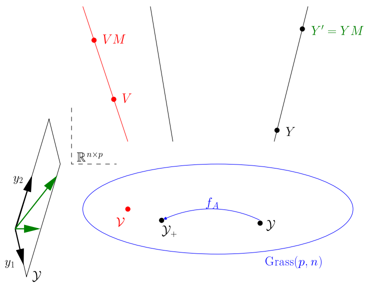
Iterations on the Grassmann manifold



Outline

- Eigenspace computation: definitions.
- Eigenspace computation: an application.
- Geometric approach.
- New numerical algorithms:
  - Newton methods.
  - Shifted inverse iterations.
  - Global behaviour and improvements.

Newton method for eigenspace computation

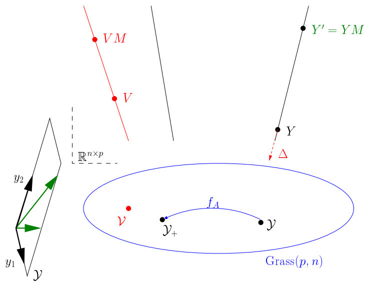


$F(Y) := \Pi_{Y_\perp} AY$, where
$\Pi_{Y_\perp} := I - Y(Y^TY)^{-1}Y^T$.

Newton method in $\mathbb{R}^{n\times p}$:

$$F(Y) + \mathrm{D}F(Y)[\Delta] = 0$$
$$Y_+ = Y + \Delta$$

Newton method for eigenspace computation



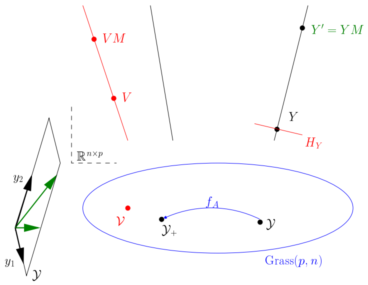
$F(Y) := \Pi_{Y_\perp} AY$, where
$\Pi_{Y_\perp} := I - Y(Y^TY)^{-1}Y^T$.

Newton method in $\mathbb{R}^{n\times p}$:

$$F(Y) + \mathrm{D}F(Y)[\Delta] = 0$$
$$Y_+ = Y + \Delta$$

Result: $\Delta = -Y$,
hence $Y_+ = 0$.

Newton method – Horizontal update



$F(Y) := \Pi_{Y_\perp} AY$, where
$\Pi_{Y_\perp} := I - Y(Y^TY)^{-1}Y^T$.

Newton method with constrained update:

$$\begin{cases} F(Y) + \mathrm{D}F(Y)[\Delta] = 0 \\ Y^T\Delta = 0 \end{cases}$$

Overdetermined system!

Newton method - Relaxations

Overdetermined system:

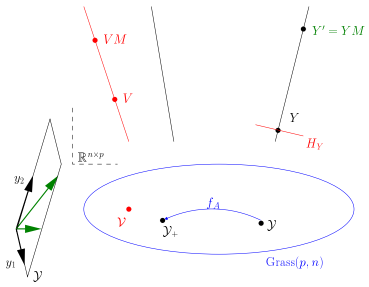
$$\begin{cases} F(Y) + \mathrm{D}F(Y)[\Delta] = 0 \\ Y^T\Delta = 0 \end{cases}$$

Two remedies:

- $$\Delta = \arg\min_{Y^T\Delta=0} \|F(Y) + \mathrm{D}F(Y)[\Delta]\|^2 = 0$$

- $$\begin{cases} \Pi_{Y_\perp}(F(Y) + \mathrm{D}F(Y)[\Delta]) = 0 \\ Y^T\Delta = 0 \end{cases}$$

Newton method – Riemannian interpretation



$F(Y) := \Pi_{Y_\perp} AY$

$$\begin{cases} \Pi_{Y_\perp}(F(Y) + \mathrm{D}F(Y)[\Delta]) = 0 \\ Y^T\Delta = 0 \end{cases}$$

$Y_+ = Y + \Delta.$

Riemannian Newton method:

$$F(Y) + \nabla_\Delta F(Y) = 0$$
$$Y_+ = \mathrm{Exp}(\Delta)$$

Newton method for eigenspace computation

Iteration mapping: $\mathrm{Grass}(p,n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in \mathrm{Grass}(p,n)$ defined by
1. Pick a basis $Y \in \mathbb{R}^{n\times p}$ that spans $\mathcal{Y}$ and solve the equation

$$\Pi A \Pi \Delta - \Delta(Y^TY)^{-1}Y^TAY = -\Pi AY \qquad (1)$$

under the constraint $Y^T\Delta = 0$, where $\Pi := I - Y(Y^TY)^{-1}Y^T$.
2. Perform the update

$$\mathcal{Y}_+ = \mathrm{span}(Y + \Delta). \qquad (2)$$

Newton method – Related work

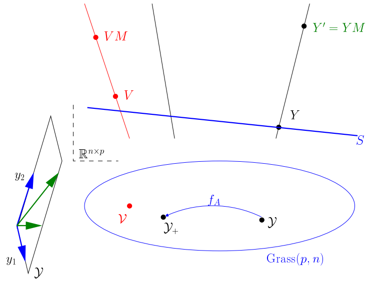
Previous work:

- Stewart [Ste73], Dongarra, Moler and Wilkinson [DMW83], Chatelin [Cha84], Demmel [Dem87].
- Lösche, Schwetlick and Timmerman [LST98].
- Edelman, Arias and Smith [EAS98].
- Lundström and Elden [LE02].

New results (PAA, Mahony, Sepulchre [AMS02]):

- Iteration converges locally quadratically (even when $A \neq A^T$).
- Convergence is locally cubic iff $A^T\mathcal{V} \subseteq \mathcal{V}$.

Newton method – Cross sections

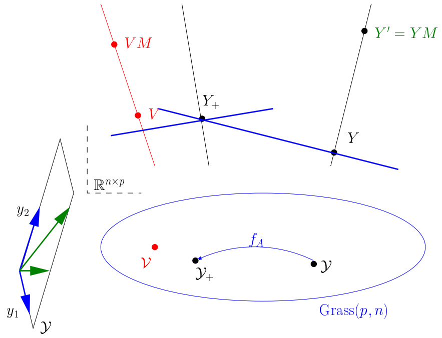


Parameterize S:
$Y = W + W_\perp K$.
Then $\Pi A Y = 0$ becomes a quadratic matrix equation

$$A_{22}K - KA_{11} + A_{21} - KA_{12}K = 0$$

where $A_{11} = W^T A W$, $A_{12} = W^T A W_\perp$, $A_{21} = W_\perp^T A W$ and $A_{22} = W_\perp^T A W_\perp$.

Newton method – Adaptive cross section

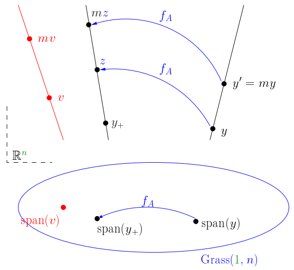


Advantage: if $A = A^T$, convergence is cubic, instead of quadratic in the fixed cross section case.

Outline

- Eigenspace computation: definitions.
- Eigenspace computation: an application.
- Geometric approach.
- New numerical algorithms:
  - Newton methods.
  - Shifted inverse iterations.
  - Global behaviour and improvements.

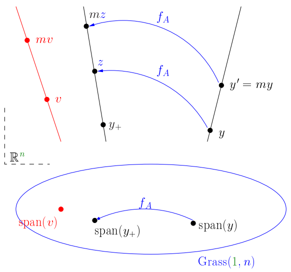
Rayleigh Quotient Iteration – Case $p = 1$



$(A - \rho(y)I)z = y$
$\rho(y) := (y^T A y)/(y^T y)$.

Local cubic convergence to eigendirections of $A = A^T$.

Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



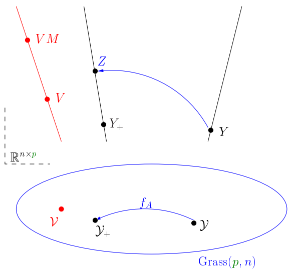
$(A - \rho(y)I)z = y$
$\rho(y) := (y^T Ay)/(y^T y).$

$\xrightarrow{p>1}$??

Want:

- Well defined on $\mathrm{Grass}(p, n)$.
- Reduces to RQI when $p = 1$.
- Local cubic convergence.

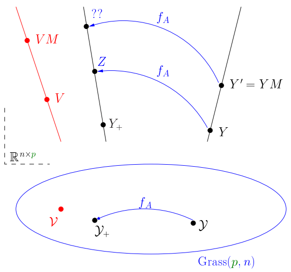
Rayleigh Quotient Iteration – Case $p \geq 1$



Classical Rayleigh Quotient Iteration ($p = 1$):
$(A - \rho(y)I)z = y$
$\rho(y) := (y^T Ay)/(y^T y).$

Generalized iteration ($p \geq 1$):
$AZ - Z(Y^T Y)^{-1} Y^T AY = Y.$

Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



$(A - \rho(y)I)z = y$
$\rho(y) := (y^T Ay)/(y^T y).$

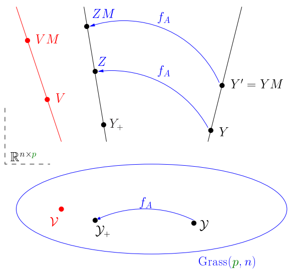
Generalized iteration:
$AZ - Z(Y^T Y)^{-1} Y^T AY = Y.$
Want:

→ Well defined on $\mathrm{Grass}(p, n)$.

→ Reduces to RQI when $p = 1$.

→ Local cubic convergence.

Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



$(A - \rho(y)I)z = y$
$\rho(y) := (y^T Ay)/(y^T y).$

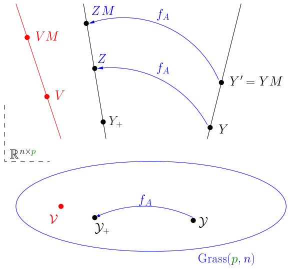
Generalized iteration:
$AZ - Z(Y^T Y)^{-1} Y^T AY = Y.$
Want:

→ Well defined on $\mathrm{Grass}(p, n)$.

→ Reduces to RQI when $p = 1$.

→ Local cubic convergence.

Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



$(A - \rho(y)I)z = y$

$\rho(y) := (y^T Ay)/(y^T y)$.

Generalized iteration:

$AZ - Z(Y^T Y)^{-1} Y^T AY = Y$.

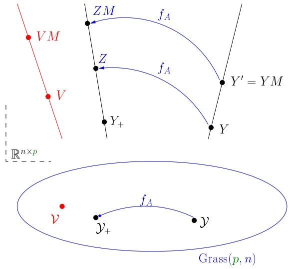
Want:

$\rightarrow$ Well defined on Grass(p, n).

$\rightarrow$ Reduces to RQI when $p = 1$.

$\rightarrow$ Local cubic convergence.

Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



$(A - \rho(y)I)z = y$

$\rho(y) := (y^T Ay)/(y^T y)$.

Generalized iteration:

$AZ - Z(Y^T Y)^{-1} Y^T AY = Y$.

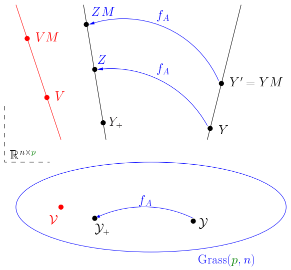
Want:

$\rightarrow$ Well defined on Grass(p, n).

$\rightarrow$ Reduces to RQI when $p = 1$.

$\rightarrow$ Local cubic convergence.

Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



$(A - \rho(y)I)z = y$

$\rho(y) := (y^T Ay)/(y^T y)$.

Generalized iteration:

$AZ - Z(Y^T Y)^{-1} Y^T AY = Y$.

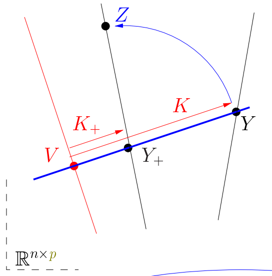
Want:

$\rightarrow$ Well defined on Grass(p, n).

$\rightarrow$ Reduces to RQI when $p = 1$.

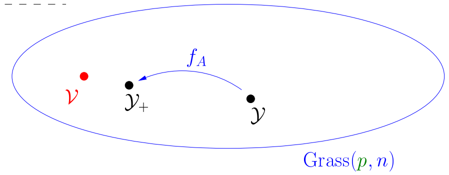
$\rightarrow$ Local cubic convergence.

Rayleigh Quotient Iteration – Local Convergence

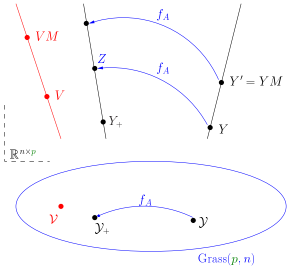


$AZ - Z(Y^T Y)^{-1} Y^T AY = Y$

Result: $\|K_+\| \leq c\|K\|^3$



Rayleigh Quotient Iteration – From $p = 1$ to $p \geq 1$



$(A - \rho(y)I)z = y$

$\rho(y) := (y^T Ay)/(y^T y)$.

Generalized iteration:

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

Want:

→ Well defined on Grass(p, n).

→ Reduces to RQI when $p = 1$.

→ Local cubic convergence.

Grassmann Rayleigh Quotient Iteration

Iteration mapping: Grass$(p, n) \ni \mathcal{Y} \mapsto \mathcal{Y}_+ \in$ Grass(p, n) defined by

1. Pick a basis $Y \in \mathbb{R}^{n\times p}$ that spans $\mathcal{Y}$.
2. Solve

$$AZ - Z(Y^TY)^{-1}Y^TAY = Y \tag{3}$$

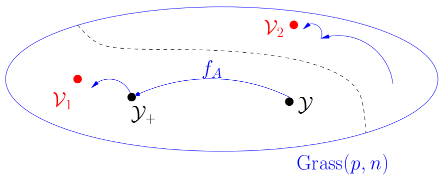
for $Z \in \mathbb{R}^{n\times p}$.

3. Define $\mathcal{Y}_+$ as the span of Z.

Outline

- Eigenspace computation: definitions.
- Eigenspace computation: an application.
- Geometric approach.
- New numerical algorithms:
  - Newton methods.
  - Shifted inverse iterations.
  - Global behaviour and improvements.

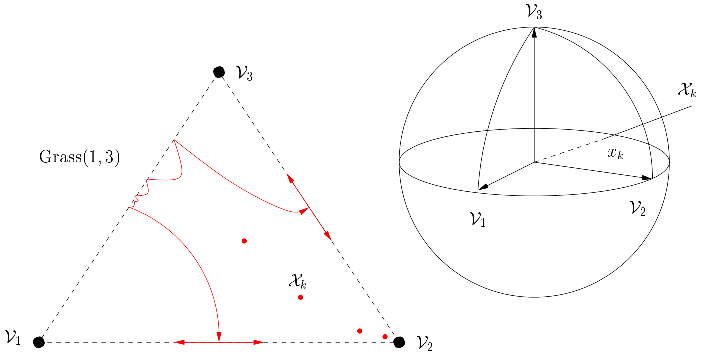
Global behaviour of iterative methods



Study basins of attraction of eigenspaces.

Global behaviour for $p = 1$

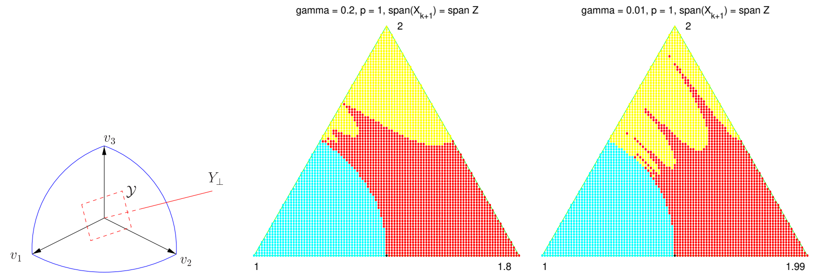
When $p = 1$,
Riemannian Newton method $\equiv$ Rayleigh Quotient Iteration.



Reference: Batterson and Smillie [BS89].

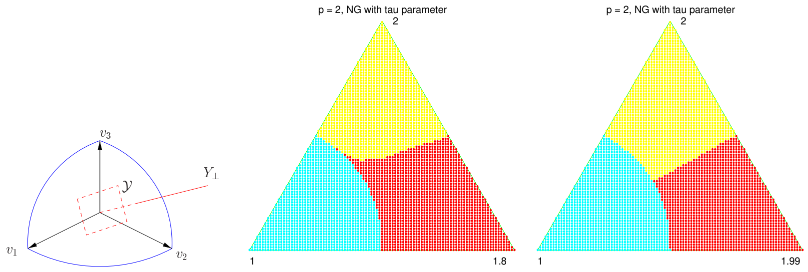
Newton method – Global behaviour ($p > 1$)

Case $p = 2$, $n = 3$:



Basins may deteriorate when eigenvalues are clustered.

Newton method – Improved global behaviour

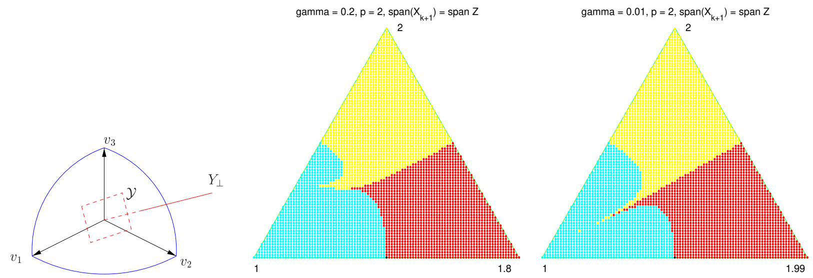


Remedy: modified Newton methods with large basins of attraction.

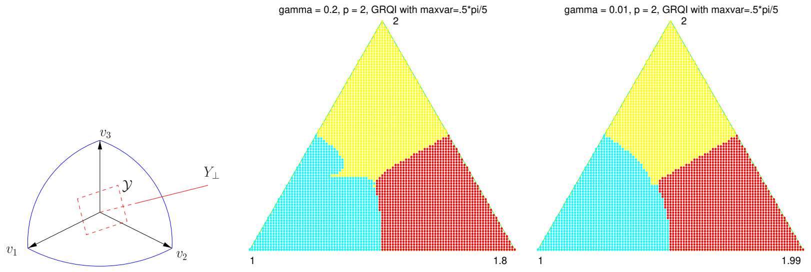
Rayleigh Quotient Iteration – Global behaviour ($p > 1$)

$$AZ - Z(Y^TY)^{-1}Y^TAY = Y, \; Y_+ = ZM$$

Case $p = 2$, $n = 3$:



Rayleigh Quotient Iteration – Improved global behaviour



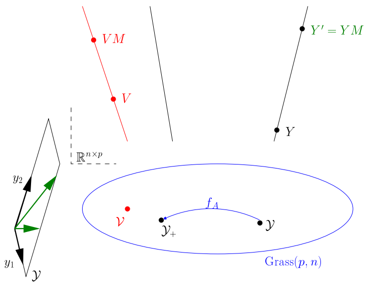
Modified Rayleigh Quotient Iteration with limited steps.

Reference: PAA, Sepulchre, Van Dooren, Mahony [ASVM03].

Refining eigenspace estimates – Main Results

- Newton methods:
  - Practical formulation of the Riemannian Newton method on the Grassmann manifold.
  - Newton method for refining eigenspace estimates. Quadratic convergence in general. Cubic convergence if $A = A^T$.
- Shifted inverse iterations:
  - Grassmannian generalization of the Rayleigh Quotient Iteration, using block shifts. Local cubic convergence to the eigenspaces of $A = A^T$.
  - Two-sided version for the nonsymmetric case.
- Study of the global behaviour of the iterations. Modified methods with large basins of attraction and high rate of convergence.

Refining eigenspace estimates – Conclusion



→ Geometry: iteration on the *Grassmann manifold* of p-planes in $\mathbb{R}^n$.

→ Dynamics: $\mathcal{Y}_{k+1} = f_A(\mathcal{Y}_k)$. Discrete dynamical system.

→ Numerical analysis: compute $f_A(\mathcal{Y}_k)$ efficiently.

References

[AMS02] P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, submitted to Acta Appl. Math., November 2002.

[ASVM03] P.-A. Absil, R. Sepulchre, P. Van Dooren, and R. Mahony, *Cubically convergent iterations for invariant subspace computation*, submitted to SIAM J. Matrix Anal. Appl., January 2003.

[BS89] S. Batterson and J. Smillie, *The dynamics of Rayleigh quotient iteration*, SIAM J. Numer. Anal. **26** (1989), no. 3, 624–636.

[Cha84] F. Chatelin, *Simultaneous Newton's iteration for the* *eigenproblem*, Computing, Suppl. **5** (1984), 67–74.

[Dem87] J. W. Demmel, *Three methods for refining estimates of invariant subspaces*, Computing **38** (1987), 43–57.

[DMW83] J. J. Dongarra, C. B. Moler, and J. H. Wilkinson, *Improving the accuracy of computed eigenvalues and eigenvectors*, SIAM J. Numer. Anal. **20** (1983), no. 1, 23–45.

[EAS98] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl. **20** (1998), no. 2, 303–353.

[LE02] E. Lundström and L. Eldén, *Adaptive eigenvalue computations using Newton's method on the Grassmann manifold*, SIAM J. Matrix Anal. Appl. **23** (2002), no. 3, 819–839.

[LST98] R. Lösche, H. Schwetlick, and G. Timmerman, *A modified block Newton iteration for approximating an invariant subspace of a symmetric matrix*, Linear Algebra Appl. **275-276** (1998), 381–400.

[Ste73] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review **15** (1973), no. 4, 727–764.

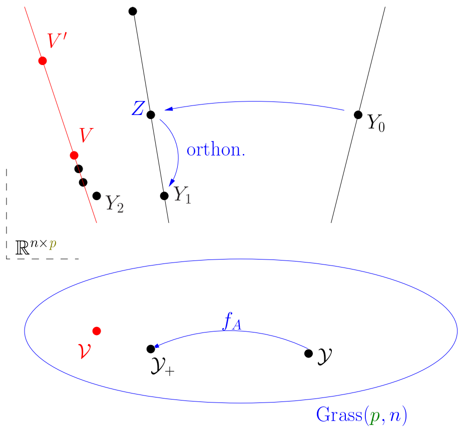
Numerical experiment:

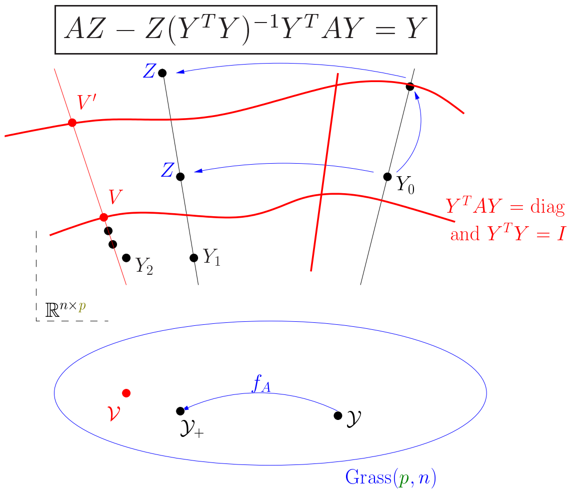
$$AZ - ZY^TAY = Y,\ Y^TY = I$$

$$Y_+ = \mathrm{qf}(Z).$$

```
A =
       1       0       0
       0       2       0
       0       0       3
Y =
 -1.0762e-001 -4.2770e-002
  9.9334e-001  3.6714e-002
 -4.1138e-002  9.9841e-001
Y =
 -1.3004e-003 -2.9822e-004
 -9.8317e-001 -1.8269e-001
  1.8269e-001 -9.8317e-001
Y =
 -4.4217e-011  2.3683e-009
  1.8740e-002 -9.9982e-001
 -9.9982e-001 -1.8740e-002
```


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




$AZ - Z(Y^TY)^{-1}Y^TAY = Y$
Z
V'
Z
Y_0
V
$Y^TAY = \mathrm{diag}$
and $Y^TY = I$
Y_2
Y_1
$\mathbb{R}^{n\times p}$
f_A
$\mathcal{V}$
$\mathcal{Y}_+$
$\mathcal{Y}$
$\mathrm{Grass}(p, n)$