Workshop on Lie group methods and control theory June 28 - July 1 Edinburgh

## Numerical techniques for approximating the solution of matrix ODE on the general linear group

Nicoletta Del Buono

Joint work with : Luciano Lopez



# Outline

- The matrix ODE we deal with
- Theoretical results
- ✤ Examples
- Numerical tools:
  - Substituting approach
  - Solution via Riccati equation
  - ➢ SVD approach
- Rectangular Case
- Numerical examples

# **The differential system**

Consider the matrix differential equation

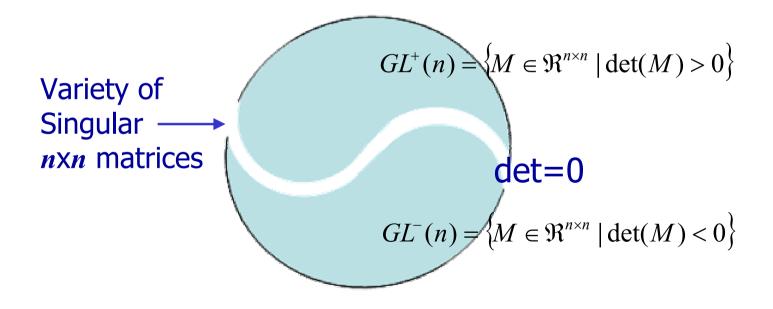
•  

$$Y(t) = Y(t)^{-T} F(Y(t), Y(t)^{-T})$$
  
 $Y(0) = Y_0 \in GL(n)$ 

- *F* is a continuous matrix function, globally Lipschitz on a subdomain of GL(n)
- the solution Y(t) exists and is unique in a neighborhood ]-τ τ[ of the origin 0

# The structure of GL(n)

Two maximal connected and disjoint open subsets comprising GL(n)



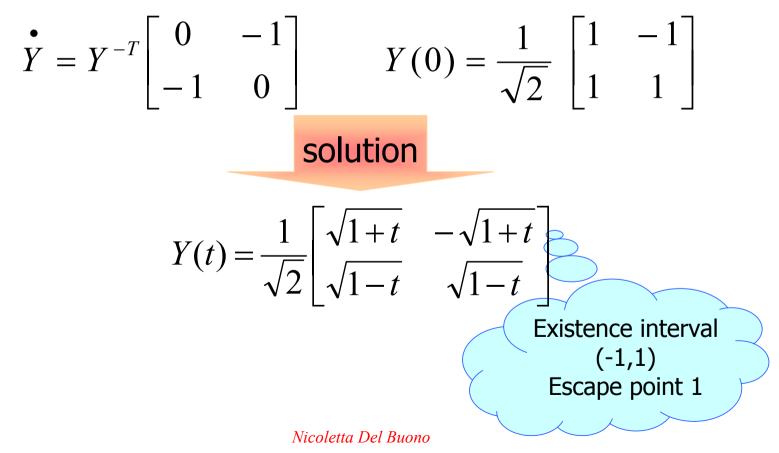
## **Theoretical results**

- The existence of the solution Y(t) for all t is not guaranteed a priori and the presence of a finite escape time behavior is not precluded.
- The value of the escape point depends on the function F
  - ➢ If the escape point *τ* is finite then *Y*(*t*) approaches a singular matrix as *t* → *τ*

 $\succ$  if *τ* < ∞ then *Y*(*t*) exists for all *t* > *θ* 

## **Theoretical results**

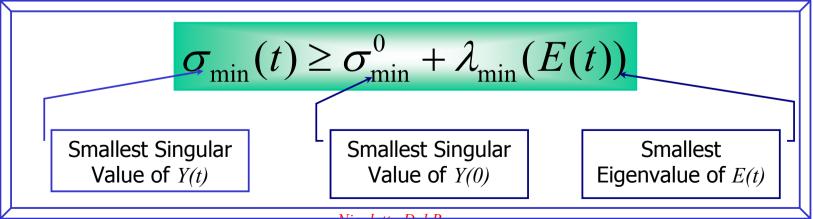
**Example:** F constant function with trace(F) = 0



## **Theoretical results**

Relationship between the singular values of the solution Y(t), the initial condition Y(0) and the symmetric matrix function:

$$E(t) = \int_{0}^{T} [F^{T}(Y(s), Y^{-T}(s)) + F(Y(s), Y^{-T}(s))] ds$$



Nicoleíía Del Buono

## **Systems with structure**

If the matrix function *F* maps all matrices into the Lie algebra of skew-symmetric matrices
 *Y(t)* belongs to the orthogonal manifold (whenever *Y(0)* is orthogonal)

• If diag(F) = 0 for all nonsingular matrices  $diag(Y(t)^TY(t)) = diag(Y(0)^TY(0))$ 

## **Examples**

- **Control Theory** 
  - Optimal system assignment via Output Feedback Control
  - Balanced Matrix Factorizations
  - Balanced realizations (Isodynamical flows)
- \* Multivariate Data Analysis
  - Weighted Oblique Procrustes problem
- **\*** Inverse Eigenvalue Problem
  - Pole placement or eigenvalue assignment problem via output feedback
  - > Prescribed Entries Inverse Eigenvalue Problem

#### Output Feedback Control of linear system

➤ Consider the linear dynamical system defined by the triple (A,B,C)∈P<sup>n×n</sup>×P<sup>n×m</sup>×P<sup>p×n</sup>

 $\begin{aligned} \mathbf{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$ 

- The process of "*feeding back*" the output or the state variables in a dynamical system configuration through the input channels
- > Output Feedback: u(t) is replaced by u(t)=Ky(t)+v(t)

 $K \in \mathbb{P}^{m \times p}$  feedback gain matrix

Nicoletta Del Buono

#### Output Feedback Control of linear system

> The feedback system is

$$\begin{aligned} \mathbf{\dot{x}}(t) &= (A + BKC) \mathbf{x}(t) + Bv(t) \\ y(t) &= Cx(t) \end{aligned}$$

#### Optimal system assignment

➢ Given a target system described by the triple
(F,G,H)∈P<sup>n×n</sup>×P<sup>n×m</sup>×P<sup>p×n</sup> find an optimal feedback transformation of (A,B,C) which results the best approximation of (F,G,H).

★ The set GL(n)×P<sup>m×p</sup> of feedback transformation is a Lie group under the operation (T<sub>1</sub>, K<sub>1</sub>)◦(T<sub>2</sub>,K<sub>2</sub>) = (T<sub>1</sub>T<sub>2</sub>, K<sub>1</sub>+K<sub>2</sub>)

We can consider action on the output feedback group and orbits, particularly:

 $\Phi(A,B,C) = \{ (T(A+BKC)T^{-1},TB,CT^{-1} | (T,K) \in GL(n) \times \mathbb{P}^{m \times p} \}$ 

★ The distance function  $\Phi = \|T(A + BKC)T^{-1} - F\|^{2} + \|TB - G\|^{2} + \|CT^{-1} - H\|^{2}$ 

 The gradient flow of this distance function with respect to a specific Riemannian metric on Φ(*A*,*B*,*C*) can be written as:

•  

$$T = T^{-T} f(T, T^{-T}, K)$$
  
•  
 $K = -B^{T}T^{T} (T(A + BKC)T^{-1} - F)T^{-T}C^{T}$ 

#### Balanced matrix factorizations

- General matrix factorization problem:
  - Given a matrix  $H \in P^{k \times l}$  find two  $X \in P^{k \times n}$  and  $Y \in P^{n \times l}$ such that H = XY
    - $\succ$  balanced factorization  $X^T X = Y Y^T$
    - > diagonal balanced factorization  $X^T X = Y Y^T = D$
- Balanced and diagonal balanced factorization can be characterized as critical points of cost functions defined on the orbit

 $O(X,Y) = \{(XT^{-1},TY) \in \mathbf{P}^{k \times n} \times \mathbf{P}^{n \times l} \mid T \in GL(n)\}$ 

The cost functions are respectively:

 $\Phi: \mathcal{O}(X, Y) \to \mathcal{P} \qquad \Phi(XT^{-1}, TY) = ||XT^{-1}||^2 + ||TY||^2$  $\Phi_N: \mathcal{O}(X, Y) \to \mathcal{P} \qquad \Phi_N(XT^{-1}, TY) = tr(NT^{-T}X^TXT^{-1} + NTYY^TT^T)$ 

Applying a gradient flow techniques differential systems on *GL(n)* can be constructed:

balanced  

$$T = T^{-T} (X^T X (T^T T)^{-1} - T^T TYY^T) \qquad T(0) = T_0$$

$$T = T^{-T} (X^T X T^{-1} N T^{-T} - T^T N TYY^T) \qquad T(0) = T_0$$
diagonal balanced

#### **\*** Balanced realizations in linear system theory

Consider the linear dynamical system defined by the triple  $(A,B,C) \in \mathbb{P}^{n \times n} \times \mathbb{P}^{n \times m} \times \mathbb{P}^{p \times n}$   $\mathbf{x}(t) = Ax(t) + Bu(t)$ y(t) = Cx(t)

Scamians:  $W_C = \int_0^\infty e^{At} B B^T e^{A^T t} dt$   $W_O = \int_0^\infty e^{A^T t} C^T C e^{At} dt$ (A,B,C) is a balanced realization if  $W_C = W_O$ 

 $\succ$  (*A*,*B*,*C*) is a **diagonal balanced realization** if  $W_C = W_O = D$ 

★ Any *T*∈*GL*(*n*) changes a realization by  $(A,B,C) \rightarrow (TAT^{-1}, TB, CT^{-1})$ 

and the Gramians via

 $W_C \rightarrow T W_C T^{-1}$ 

 $W_0 \rightarrow T^{-T} W_0 T^{-1}$ 

Balanced and diagonal balanced realizations have been proved to be critical points of costs functions defined on the orbit

 $O(A, B, C) = \{ (TAT^{-1}, TB, CT^{-1}) \in P^{n \times n} \times P^{n \times m} \times P^{k \times n} | T \in GL(n) \}$ 

The cost functions are respectively:

 $\Phi: \mathcal{O}(A, B, C) \to \mathcal{P} \qquad \Phi(T) = tr(TW_C T^{-1} + T^{-T}W_O T^{-1})$  $\Phi_N: \mathcal{O}(A, B, C) \to \mathcal{P} \qquad \Phi_N(T) = tr(NTW_C T^{-1} + NT^{-T}W_O T^{-1})$ 

All balancing transformation *T*∈*GL*(*n*) for a given asymptotically stable system (*A*,*B*,*C*) can be obtained solving the gradient flow

balanced 
$$\overset{\bullet}{T} = T^{-T} (W_O (T^T T)^{-1} - T^T T W_C)$$
  $T(0) = T_0$   
 $\overset{\bullet}{T} = T^{-T} (W_O T^{-1} N T^{-T} - T^T N T W_C)$   $T(0) = T_0$  diagonal balanced

Nicoletta Del Buono

## Examples in Multivariate Data Analysis

Weighted oblique Procrustes problem (WObPP)

> Manifold of the oblique rotation matrices

 $OB(n) = \{X \in \mathbb{P}^{n \times n} \mid \det(X) \neq 0, diag(X^T X) = I\}$ 

Given A,B,C fixed matrices with conformal dimensions

> Minimize || AXC- B || subject to  $X \in OB(n)$ 

Problem in factor analysis known as a "rotation to factor-structure matrix"

> Minimize  $|| AX^{-T}C^{-}B ||$  subject to  $X \in OB(n)$ 

Problem of finding an approximation to a "factorpattern" matrix

## Examples in Multivariate Data Analysis

The solution of the WObPP problem can be obtained solving a descent matrix ODE:

$$\frac{dX}{dt} = -\pi_{OB(n)}(\nabla) = -X^{-T}off(X^{T}\nabla)$$

♦ being  $\nabla$  the gradient of the function to be minimize with respect to the chosen metric

(N. Trendafilov FGCS 2003)

## Examples in Inverse Eigenvalue Problem and control theory

# Pole placement or eigenvalue assignment via output feedback:

- Siven a linear system described by the triple (A,B,C)and a self-conjugate set of complex points  $\{\lambda_1 \lambda_2 \dots \lambda_{n_n}\}$
- ➢ find a feedback gain matrix K such that A+BKC has eigenvalues  $λ_i$

◆ Denoted by Λ a fixed matrix with eigenvalues λ<sub>i</sub> the pole placement task is equivalent to find a matrix *T*∈*GL*(*n*) and *K*∈P<sup>*m×p*</sup> minimizing the distance μι ||Λ−*T*(*A*+*BKC*)*T*<sup>-1</sup>||

## Examples in Inverse Eigenvalue Problem and control theory

Using a gradient flow techniques the solution can be obtained solving

•  

$$T = T^{-T} [(A + BKC)^T, T^T (\Lambda - (A + BKC))T^{-T}]$$
  
•  
 $K = -B^T T^T (T (A + BKC)T^{-1} - F)T^{-T}C^T$ 

### Examples in Inverse Eigenvalue Problem

Matrix completion with prescribed eigenvalues

# PEIEP (prescribed entries inverse eigenvalue problem) :

Given  $\lambda = \{(i_v, j_v) \mid v = 1, ..., m\} \text{ } m \text{ pairs of integers } 1 \le i_v < j_v \le n$   $\lambda = \{a_1, ..., a_m\} \subset P$   $\lambda_1, ..., \lambda_n \} \subset X \text{ closed under conjugation}$ Find a matrix  $X \in P^{n \times n}$  such that  $\sigma(X) = \{\lambda_1, ..., \lambda_n\}$ and  $x_{i_v, j_v} = a_v, v = 1, ..., m$ 

### Examples in Inverse Eigenvalue Problem

★ Let  $\Lambda$  a matrix with eigenvalues  $\lambda_i$  and denoting  $M(\Lambda) = \{V\Lambda V^{-1} \mid V \in GL(n)\}$ 

the orbit of matrices isospectral to  $\Lambda$  under the action group of GL(n) and

$$\Sigma(\Lambda, \boldsymbol{a}) = \{ X = [x_{ij}] \in \mathbb{P}^{n \times n} \mid x_{i_v j_v} = a_v \quad v = 1, \dots, m \}$$

\* Solving the PEIEP is to find intersection of the two geometric entities  $M(\Lambda)$  and  $\Sigma(\Lambda, a)$ 

#### Examples in Inverse Eigenvalue Problem

★ Minimize for each given *X* ∈ M(Λ) the distance between *X* and Σ(Λ,*a*)

$$\min_{V \in M(\Lambda)} \frac{1}{2} < V\Lambda V^{-1} - P(V\Lambda V^{-1}), V\Lambda V^{-1} - P(V\Lambda V^{-1}) >$$
  
Projection on  $\Sigma(\Lambda, a)$ 

Using a descent flow approach we get

 $\frac{dV}{dt} = \kappa (V\Lambda V^{-1})V^{-T} \quad \text{with} \quad \kappa(X) = [X^T, X - P(X)]$ 

(M.T. Chu et al. FGCS 2003)

### Numerical Approximation: substituting approach

Consider our system:

•  

$$Y(t) = Y(t)^{-T} F(Y(t), Y(t)^{-T})$$
  
 $Y(0) = Y_0 ∈ GL(n)$ 

Setting  $Z=Y^{-T}$  from  $Y^{T}Z=I$  we get

$$Y^{T} Z + Y^{T} \dot{Z} = 0 \Leftrightarrow \dot{Z} = -Y^{-T} Y^{T} Z$$

$$Y = ZF(Y,Z) = H(Y,Z), \qquad Y(0) = Y_0$$

•  

$$Z = -ZF^{T}(Y,Z)Z^{T}Z = -ZH^{T}(Y,Z)Z, \quad Z(0) = Y_{0}^{-T}$$

Nicoletta Del Buono

# **Substituting Approach**

#### Advantages:

> No direct use of the inverse of Y(t) (computational advantages)

#### Drawbacks:

- Solution of a new matrix ODE with double dimension with respect to the original system;
- High stiffness (when Y(t) tends to a singular matrix or the Lipschitz constant of H is large);
- The presence of an additional structure of the solution matrix Y(t) is not considered need of ad hoc numerical scheme

• When the matrix function F does not depend explicitly on  $Y^{-T}$ , i.e.:

$$\overset{\bullet}{Y(t)} = Y(t)^{-T} F(Y(t))$$
$$Y(0) = Y_0 \in GL(n)$$

It could be convenient work with the implicit equation

$$Y(t)Y(t) = F(Y(t))$$
$$Y(0) = Y_0 \in GL(n)$$

Nicoletta Del Buono

Applying the second order Gauss Legendre method, we get:

$$Y_{n+1}^{T}Y_{n+1} + Y_{n}^{T}Y_{n+1} - Y_{n+1}^{T}Y_{n} - Y_{n}^{T}Y_{n} - 2hF\left(\frac{Y_{n} + Y_{n+1}}{2}\right) = 0$$

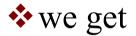
The previous equation can be iteratively solved starting from an initial approximation  $Y_{n+1}^{(0)}$ 

(avoiding the nonlinearity of F)

$$Y_{n+1}^{T}Y_{n+1} + Y_{n}^{T}Y_{n+1} - Y_{n+1}^{T}Y_{n} - Y_{n}^{T}Y_{n} - 2hF\left(\frac{Y_{n} + Y_{n+1}^{(0)}}{2}\right) = 0$$

The latter equation is the prototype of an Algebraic Riccati equation, in fact setting

$$A = Y_n$$
 and  $C = Y_n^T Y_n + 2hF\left(\frac{Y_n + Y_{n+1}^{(0)}}{2}\right)$ 



$$R(X) = X^T X + A^T X - X^T A + C = 0$$

## Solution via Algebraic Riccati equation

- Numerical methods to solve Algebraic Riccati equation are based on fixed point or Newton iteration:
  - > Picard iteration:

$$A^{T}X_{k+1} - X_{k+1}^{T}A = -C - X_{k}^{T}X_{k}$$

- > Newton method:
  - $\succ R: \mathbf{P}^{n \times n} \to \mathbf{P}^{n \times n}$

> its Frechét derivatitive is:  $R'_X(H) = H^T(X - A) + (X + A)^T H$ 

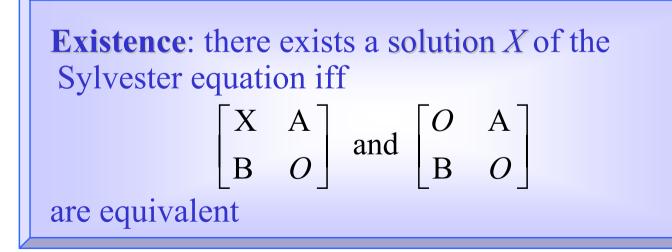
➢ the Newton iteration starts from X<sub>θ</sub> and solves R(X)=0 via  $X_{k+1}=X_k+D_k$  being  $D_k$  the solution of Sylvester equation

 $R'_{X}(D_{k}) = -R(X_{k}) \Leftrightarrow (X_{k} + A)^{T} D_{k} + D_{k}^{T} (X_{k} - A) = -R(X_{k})$ 

Solving Riccati equation implies the numerical treatment of the Sylvester equation

 $\mathbf{A}X + X^{\mathrm{T}}\mathbf{B} = \mathbf{X}$ 

with A, B, X given  $n \times n$  matrices



✤ To obtain conditions for uniqueness of solution and for constructing it, we reformulate the Sylvester equation as a  $n^2 \times n^2$  linear system:

 $(I \otimes A) \operatorname{vec}(X) + (B^T \otimes I) \operatorname{vec}(X^T) = \operatorname{vec}(X)$ 

$$\operatorname{vec}(X^{T}) = P(n,n)\operatorname{vec}(X)$$
$$P(n,n) = \sum_{i=1}^{n} \sum_{j=1}^{n} E_{ij} \otimes E_{ij}^{T}$$

$$\underbrace{\left[(I \otimes A) + (B^T \otimes I)P(n,n)\right]}_{M} \operatorname{vec}(X) = \operatorname{vec}(X)$$

Nicoletta Del Buono

$$M = \begin{bmatrix} A + e_1 b_1^T & e_2 b_1^T & \cdots & e_n b_1^T \\ e_1 b_2^T & A + e_2 b_2^T & \cdots & e_n b_2^T \\ \vdots & \vdots & \ddots & \vdots \\ e_1 b_n^T & e_2 b_n^T & \cdots & A + e_n b_n^T \end{bmatrix}$$

being  $b_i$  the columns of the matrix B

**Uniqueness:** there exists a unique solution *X* of the Sylvester equation  $AX+X^TB=X$  if the matrix *M* is non-singular (rank  $(M)=n^2$ )

Nicoletta Del Buono

Considering the linear equation derived from:

- **Picard iteration:**  $A = A^T$  and  $B = A \Rightarrow M$  is singular
- ▶ Newton iteration:  $A = X_k + A^T$  and  $B = X_k A \implies M$  is non-singular ⇒ unique solution !
- Newton method converges in a reasonable number of iterations
- Numerical solution of Sylvester equation :
  - Direct methods (QR, Gaussian Elimination);
  - Iterative algorithms;
  - ➢ Generalize Conjugate Residual method.

## **Singular Value Decomposition**

- To avoid the inverse matrix computations and to control the singularities of the matrix solution Y(t) we can adopt a continuous Singular Value Decomposition approach
- \* The continuous SVD of Y(t) is a continuous factorization

 $Y(t) = U(t) \Sigma(t) V^{\mathrm{T}}(t)$ 

- $\succ U(t), V(t)$  orthogonal matrices  $(U^T U = I_n \text{ and } V^T V = I_n)$
- >  $\Sigma(t)$  diagonal matrix with diagonal elements the singular values  $\sigma_i(t)$  of Y(t)
- \* The motion of Y(t) is now described by the variables U(t),  $\Sigma(t)$ , V(t) giving more information on the flow

### **Singular Value Decomposition**

Suppose that the solution Y(t) possesses dinstinct and nonzero singular values σ<sub>i</sub>(t), for i=1,..., n and t in [0, τ) then there exists a continuous SVD of Y(t) and the factors U(t), Σ(t), V(t) of such a decomposition satisfy the following ODEs:

$$\begin{split} \dot{\Sigma} &= \Sigma^{-1} V^T F(Y, Y^{-T}) V - H \Sigma + \Sigma K, \quad \Sigma(0) = \Sigma_0 \\ \dot{U} &= U H, \quad U(0) = U_0 \\ \dot{V} &= V K, \quad V(0) = V_0 \end{split}$$

### **Singular Value Decomposition**

The differential equations for the singular values are

$$\overset{\bullet}{\sigma}_{i} = \frac{1}{\sigma_{i}} \left( V^{T} F(Y, Y^{-T}) V \right)_{ii}, \quad i = 1, \cdots, n$$

The elements of the skew-symmetric matrices *H*, *K* are

$$H_{ij} = \frac{1}{\sigma_i \sigma_j (\sigma_j^2 - \sigma_i^2)} \left[ \sigma_j^2 (V^T F V)_{ij} + \sigma_i^2 (V^T F V)_{ji} \right]$$
$$K_{ij} = \frac{1}{(\sigma_j^2 - \sigma_i^2)} \left[ (V^T F V)_{ij} + (V^T F V)_{ji} \right]$$

## **Singular Value Decomposition**

Numerical solution of:

- > a diagonal system in  $\sigma_i$  (information on the conditioning of the matrix solution Y(t))
- $\succ$  two linear systems in  $H_{ij}$   $K_{ij}$
- $\succ$  two orthogonal systems in U and V
  - → our aim is to preserve the non-singular behavior of the numerical solution → explicit integration of the systems in U and V (orthogonality preserved up to the order of the method)
- Drawback distinct singular values
   Block Continuous SVD

## **Rectangular case**

Some of the previous results can be extended to differential problems on the manifold

 $GL(m, n) = \{ Y \in P^{m \times n} \mid \operatorname{rank}(Y) = n \}, \quad n \le m$ 

Differential systems on *GL(m,n)* have the following form:

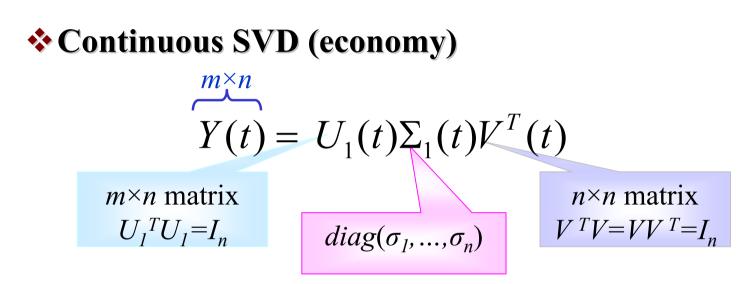
$$\mathbf{Y} = G(Y), \quad Y(0) = Y_0 \in GL(n, p)$$

\* with *G* belonging to the tangent space of GL(m,n):

$$G(Y) = Y\left(Y^{T}Y\right)^{-1}F_{1}(Y) + \left[I_{n} - Y\left(Y^{T}Y\right)^{-1}Y^{T}F_{2}(Y)\right]$$

Nicoletta Del Buono

### Rectangular Case: numerical treatment



Differentiating we obtain the differential systems satisfied by the three factors:

### Rectangular Case: numerical treatment

$$\overset{\bullet}{\sigma}_{i} = \frac{1}{\sigma_{i}} \left( V^{T} F_{1}(Y) V \right)_{ij} \qquad i = 1, \cdots, n$$
  
$$\overset{\bullet}{V} = VK, \qquad V(0) = V_{0}$$
  
$$\overset{\bullet}{U}_{1} = U_{1}H + (I_{n} - U_{1}U_{1}^{T}T)F_{2}(Y)\Sigma_{1}^{-1}, \qquad U(0) = U_{0}$$

Differential System on the Stiefel manifold

$$H_{ij} = \frac{1}{\sigma_i \sigma_j (\sigma_j^2 - \sigma_i^2)} \Big[ \sigma_j^2 (V^T F_1(Y) V)_{ij} + \sigma_i^2 (V^T F_1(Y) V)_{ji} \Big]$$
  
$$K_{ij} = \frac{1}{(\sigma_j^2 - \sigma_i^2)} \Big[ (V^T F_1(Y) V)_{ij} + (V^T F_1(Y) V)_{ji} \Big]$$

Nicoletta Del Buono

### Rectangular Case: numerical treatment

#### Substituting approach:

$$\overset{\bullet}{Y} = Y \left( Y^T Y \right)^{-1} F_1(Y) + \left[ I_n - Y \left( Y^T Y \right)^{-1} Y^T \right] F_2(Y)$$

Setting  $Z = (Y^T Y)^{-1}$  we obtain

$$\dot{Y} = YZF_{1}(Y) + \left[I - YZY^{T}\right]F_{2}(Y)$$
$$\dot{Z} = -Z\left[F_{1}(Y) + F_{1}^{T}(Y)\right]Z$$

First example:  

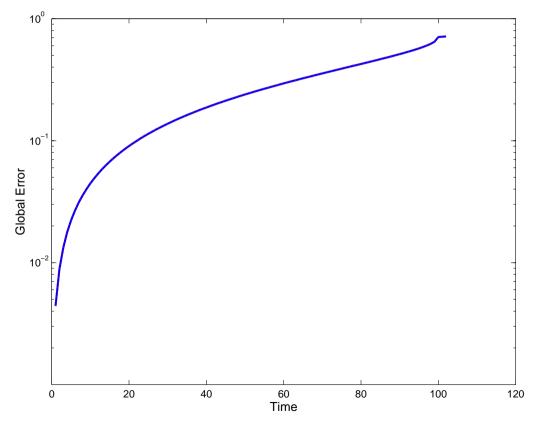
$$\dot{Y} = Y^{-T} \begin{bmatrix} 0 & -\frac{\delta}{2} \\ -\frac{\delta}{2} & 0 \end{bmatrix}$$
 $Y(0) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ 

• With solution existing in  $(-1/\delta, 1/\delta)$ 

$$Y(t) = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{1+\delta t} & -\sqrt{1+\delta t} \\ \sqrt{1-\delta t} & \sqrt{1-\delta t} \end{bmatrix}$$

• We solve the problem with  $\delta = 1/2$ 

◆ Behaviour of the global error on [0 2)



Nicoletta Del Buono

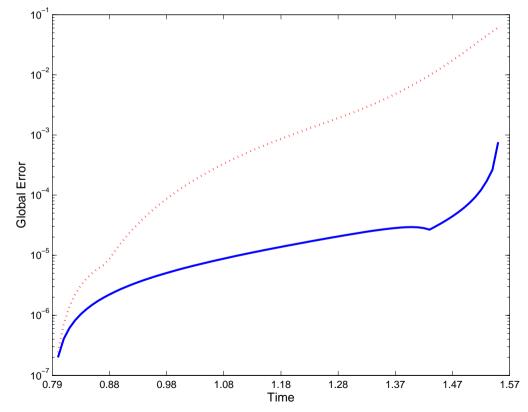
Second example

with solution

$$Y(t) = \begin{bmatrix} \cos(t) & t \\ 0 & 1 \end{bmatrix}$$

• periodically singular (for each  $\tau_k = k \pi/2$ )

Semilog plot of the global error on  $(\pi/4,\pi/2)$ 



# Conclusions

- We have considered a particular ODEs on *GL(n)* often occurring in applications
- Several problems modeled by such ODEs
- Different numerical approaches avoiding the direct use of matrix inversion and detection of singular behavior

### Future works:

Improving the validation of the proposed approaches by tackling numerical tests on real examples