Applying Stabilization Techniques to Orthogonal Gradient Flows

C. Mastroserio[1] and T. Politi[2]

[1] Dipartimento Interuniversitario di Matematica, Università di Bari
Via Orabona 4, I-70125 Bari (Italy). carmen@dm.uniba.it
[2] Dipartimento Interuniversitario di Matematica, Politecnico di Bari,
Via Orabona 4, I-70125 Bari (Italy). pptt@dm.uniba.it

Abstract. The solution of ordinary differential systems on manifolds could be treated as differential algebraic equation. In this paper we consider the solution of orthogonal differential systems deriving from the application of the gradient flow techniques to minimization problems. Neglecting the constraints for the solution a differential system is derived. Hence the problem is modified introducing a stabilization technique which is a function of the constrain. The advantage of this approach is that it is possible to apply non conservative numerical methods which are cheaper. Some numerical examples are shown.

1 Introduction

Many problem of practical interest can be modeled by systems of differential equations whose solutions satisfy some invariants, usually defined explicitely by algebraic constraints. In recent years particular attention has been paid to the development of numerical methods which approximate the solution of such a system while preserving the invariant to machine precision. These methods usually need the computational of matrix exponential once (and often repeatedly) or the solution of linear systems at each time step (see [9,6,8]) and this highly increases their computational costs.

In this paper we consider a class of differential systems derived - via a gradient flow technique - from a constraint minimization problem (on the orthogonal manifold) of a particular objective function. In this class of problems is not crucial to preserve the orthogonality of the solution, but the main interest is to get, as soon as possible, the minimum point.

Hence, we will show how it is possible to solve this differential system with invariants applying explicit methods with a splitting technique. To do this, we review the regularization technique described in [1,2] and applied to the Stiefel manifold in [4].

In [2] an important difference between the stabilization and the regularization techniques for differential system is pointed out. In regularization methods the problem is perturbed in order to obtain another system easier to solve. In this case, the solution we obtain is not the same of the initial system, hence this

P.M.A. Sloot et al. (Eds.): ICCS 2003, LNCS 2658, pp. 149–157, 2003.

implies that the perturbation to introduce must be small. On the other hand, in the stabilization techniques, the solution of the two systems (the given problem and the perturbed one) is the same, since the perturbation term vanishes when a function satisfies the constraint. Hence, the perturbation parameter does not need to be small (or large).

The paper is organized as follows: in Section 2 some stabilization techniques are recalled, in Section 3 the problem to be solved is stated and a modified gradient flow is derived introducing a perturbation term on the minimization function, in Section 4 a method for the perturbed orthogonal flow is described and finally in Section 5 some numerical tests are shown.

2 A Survey on Stabilization Techniques

Let us consider the differential system

$$Y'(t) = F(t, Y(t)), \tag{1}$$

with initial condition $Y(0) = Y_0$. For sake of simplicity we assume that there is only a unique function $Y(t)$ satisfying (1). Together with (1) we suppose that there is an invariant set (or a constraint) defined by an algebraic equation:

$$H(Y) = 0 \tag{2}$$

such that if $H(Y_0) = 0$ then $H(Y(t)) = 0$ for all $t > 0$. It is not very important to distinguish between the vector or the matrix case (i.e. Y_0 is a square real matrix and so is $Y(t)$). The important question is how to design a numerical method which preserves the properties given by the invariant set or to exploit the information on the solution given by (2) to improve the quality of the approximate solution in terms, for instance, of stability or of global error. Applications of (1)-(2) include some mechanical systems with quadratic constraints, preserving the orthogonality of a solution matrix and isospectral flows, preserving the set of the eigenvalues of the initial condition matrix. One of the most popular techniques for the stabilization of invariants is the one described by Baumgarte [3] applied to a Differential Algebraic Equation

$$\dot{y}(t) = F(y) - B(y)x$$

with invariant

$$0 = g(y).$$

Let $G = g_y(y)$ be a full rank matrix, the stabilization technique consists in replacing the invariant with a linear combination of its derivatives

$$\ddot{g}(y) + \gamma_1 \dot{g}(y) + \gamma_2 g(y) = 0$$

or

$$\dot{g}(y) + \gamma g(y) = 0.$$

In this case the approach is equivalent to replace the differential system by

$$\dot{y} = F(y) - \gamma B(GB)^{-1} g(y).$$

In [2] the following result is proved.

Theorem 1. *Consider the differential system (1) and the invariant defined by (2) and apply the stabilization*

$$Y'(t) = F(Y(t)) - \gamma F_Y(Y) H(Y) \tag{3}$$

where $F_Y = D(H_Y D)^{-1}$, *if there exists a constant* γ_0 *such that*

$$\|H_Y(Y)F(Y)\|_2 \leq \gamma_0 \|H(Y)\|_2$$

for all Y *belonging to a neighborhood of the invariant set, then it is asymptotically stable invariant manifold of (3) for* $\gamma > \gamma_0$. *In particular if* H *is an integral invariant of (1), i.e.*

$$H_Y(Y)F(Y) = 0, \qquad \forall Y,$$

then the invariant set is an asymptotically stable invariant manifold of (3) for any $\gamma > 0$.

It should be pointed out that the numerical solution can be now obtained discretizing (3) using also nonstiff integrators. This approach seems to be better than the Baumgarte and the projected invariants methods.
If we consider the differential system

$$Y'(t) = F(t, Y)Y, \qquad Y(0) = Y_0 \tag{4}$$

with $F(t, Y)$ continuous skew-symmetric function for all $Y \in \mathcal{O}_n(\mathbb{R})$ and Y_0 orthogonal square matrix of order n, it is well known that the solution $Y(t)$ is an orthogonal matrix for all $t \geq 0$. The system (4) can be considered as a system on the set of square real matrices with the nonlinear constrain

$$H(Y) = Y(t)Y^T(t) - I_n = 0 \tag{5}$$

where I_n is the identity matrix of order n. In [4] the following modified equation is proposed in order to make the manifold of orthogonal matrices attractive for the system (4):

$$Y'(t) = F(t, Y)Y - \gamma Y(t)P(Y) \tag{6}$$

where γ is a real positive number.

3 Gradient Flow Approach for the Regularization

In this section we consider the class of differential systems on the manifold of orthogonal matrices and we apply a gradient flow technique in order to minimize a given matrix function. Consider the isospectral manifold

$$M(\Lambda) = \{A \in \mathbb{R}^{n \times n} \mid A = Q^T \Lambda Q, Q \in \mathcal{O}(n)\}$$

where Λ is a given diagonal real matrix and $\mathcal{O}(n)$ is the set of $n \times n$ orthogonal matrices. The problem we are interesting in can be formulated as follows:

$$\begin{aligned} &\text{Minimize } \frac{1}{2}\|Q^T\Lambda Q - P(Q^T\Lambda Q)\|_F^2 \\ &\text{subject to } Q^TQ = I \end{aligned} \tag{7}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix and $P(\cdot)$ is the projection function. Introducing the Frobenius inner product of two matrices $A, B \in \mathbb{R}^{n\times n}$, defined by

$$\langle A, B\rangle = \text{trace}(AB^T) = \sum_{i,j=1}^{n} a_{ij}b_{ij},$$

the minimization problem (7) becomes:

$$\begin{aligned} &\text{Minimize } \frac{1}{2}\langle f(Q), f(Q)\rangle \\ &\text{subject to } g(Q) = 0 \end{aligned} \tag{8}$$

where,

$$f(Q) = Q^T\Lambda Q - P(Q^T\Lambda Q), \qquad g(Q) = Q^TQ - I.$$

Introducing a real nonnegative parameter ε it is possible to transform (8) into an unconstrained minimization problem:

$$\text{Minimize } \frac{1}{2}\langle f(Q), f(Q)\rangle + \frac{\varepsilon}{2}\langle g(Q), g(Q)\rangle. \tag{9}$$

Again we put

$$\varphi(Q,\varepsilon) = \frac{1}{2}\left[\langle f(Q), f(Q)\rangle + \varepsilon\langle g(Q), g(Q)\rangle\right]$$

and compute the Fréchet derivative of φ at Q acting on H:

$$\begin{aligned} \varphi'(Q,\varepsilon)H = \langle f'(Q)H, f(Q)\rangle + \varepsilon\langle g'(Q)H, g(Q)\rangle = \\ &= \langle H^T\Lambda Q + Q^T\Lambda H - P'(Q^T\Lambda Q)(H^T\Lambda Q + Q^T\Lambda H), f(Q)\rangle + \\ &\quad +\varepsilon\langle H^TQ + Q^TH, g(Q)\rangle = \\ &= \langle H^T\Lambda Q + Q^T\Lambda H - P(H^T\Lambda Q + Q^T\Lambda H), f(Q)\rangle + \\ &\quad +\varepsilon\langle H^TQ + Q^TH, g(Q)\rangle. \end{aligned}$$

Since the projection $P(\cdot)$ is orthogonal to function $f(Q)$, we get

$$\begin{aligned}
\varphi'(Q,\varepsilon)H &= \langle H^T \Lambda Q + Q^T \Lambda H, f(Q)\rangle + \varepsilon\langle H^T Q + Q^T H, g(Q)\rangle = \\
&= \langle H^T \Lambda Q, f(Q)\rangle + \langle Q^T \Lambda H, f(Q)\rangle + \varepsilon\langle H^T Q, g(Q)\rangle + \\
&\quad + \varepsilon\langle Q^T H, g(Q)\rangle = \\
&= \langle H, \Lambda Q f(Q)^T\rangle + \langle H, \Lambda Q f(Q)\rangle + \varepsilon\langle H, Q g(Q)^T\rangle + \varepsilon\langle H, Q g(Q)\rangle = \\
&= \langle H, \Lambda Q f(Q)^T + \Lambda Q f(Q) + \varepsilon\left(Q g(Q)^T + Q g(Q)\right)\rangle.
\end{aligned}$$

This last equation suggests that the gradient of φ at a general matrix H, with respect to the Frobenius inner product, can be interpreted as the matrix:

$$\begin{aligned}
\varphi'(Q,\varepsilon) &= \Lambda Q f(Q)^T + \Lambda Q f(Q) + \varepsilon\left(Q g(Q)^T + Q g(Q)\right) = \\
&= \Lambda Q\left(f(Q)^T + f(Q)\right) + \varepsilon Q\left(g(Q)^T + g(Q)\right) = \\
&= \Lambda Q\left(2Q^T \Lambda Q - 2P(Q^T \Lambda Q)\right) + \varepsilon Q\left(2Q^T Q - 2I\right) = \\
&= 2\Lambda Q\left[Q^T \Lambda Q - P(Q^T \Lambda Q)\right] + 2\varepsilon Q(Q^T Q - I).
\end{aligned}$$

The choice of the projection function $P(\cdot)$ leads to different classes of problem on the isospectral manifold $M(\Lambda)$. It should be observed that the obtained unconstrained differential systems obtained considering the gradient of function φ can be seen as a stabilized differential system (it is easy to observe that the last term is just the stabilized term introduced in the previous section).

4 A Numerical Method for Regularized System

Once the regularization term has been added to the given differential system to solve it some numerical aspects are to be taken into account. A possible technique is the post-stabilization of the numerical solution (see [2]). It consists in the application of a stabilization step with respect to the manifold at the end of each time step. In the case of orthogonal differential systems we can consider the following method introduced in [2] and used in [4] to solve differential systems on the Stiefel manifold in order to compute a subset of Lyapunov exponents of a dynamical system. First the system (4) (or the system (6) choosing $\gamma = 0$) is integrated from t_k to t_{k+1} using an explicit method and obtaining an approximate solution $\tilde{Y}_{k+1}$. Then the regularization term

$$Y'(t) = -\gamma Y(t) P(Y) \tag{10}$$

is integrated using, for instance, forward Euler with the same stepsize h but with γ chosen so that $\gamma h = \frac{1}{2}$ and taking as approximation at the previous step $\tilde{Y}_{k+1}$:

$$Y_{k+1} = \tilde{Y}_{k+1} - \gamma h \tilde{Y}_{k+1} P(\tilde{Y}_{k+1}) \tag{11}$$

and then

$$Y_{k+1} = \tilde{Y}_{k+1} \left(I_n - \frac{1}{2} P(\tilde{Y}_{k+1}) \right).$$

This last step can be seen as the application of one step of the Schulz method to compute the polar decomposition (see [7]).

5 Numerical Tests

Example 1. Let us consider the following gradient flow introduced in [5]. If L is a real symmetric matrix then the orthogonal flow defined by the differential system:

$$Y'(t) = Y[Y^T LY, P(Y^T LY)], \qquad Y(0) = Y_0 \tag{12}$$

with Y_0 a random orthogonal matrix, converges to the eigenvector matrix related to L. In this case the projection function is

$$P(X) = \mathrm{diag}(X)$$

while the objective function to minimize is (7). We have solved (12) using the numerical method introduced in the previous section. Figure 1 shows the orthogonal error of the numerical approximation Y_n, given by

$$E_n = \|Y_n^T Y_n - I\|_F$$

while in Figure 2 are reported the values of objective function. In Figure 3 is shown the orthogonal error using the numerical method but applying two iterates of (11). In this case the orthogonal error is smaller but considering the values of the objective function they are the same shown in Figure 2. In this case it seems that a better integration in the manifolds does not imply a speeder descent toward the minimum point of the objective function.

Example 2. Given a real symmetric matrix A we want to find a least squares approximation of A that is still symmetric but has a prescribed set of eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. In this case the orthogonal flow defined by the differential system:

$$Y'(t) = Y[Y^T LY, A], \qquad Y(0) = Y_0 \tag{13}$$

with Y_0 a random orthogonal matrix. Figure 1 shows the orthogonal error of the numerical approximation while in Figure 2 are sketched the values of objective function.

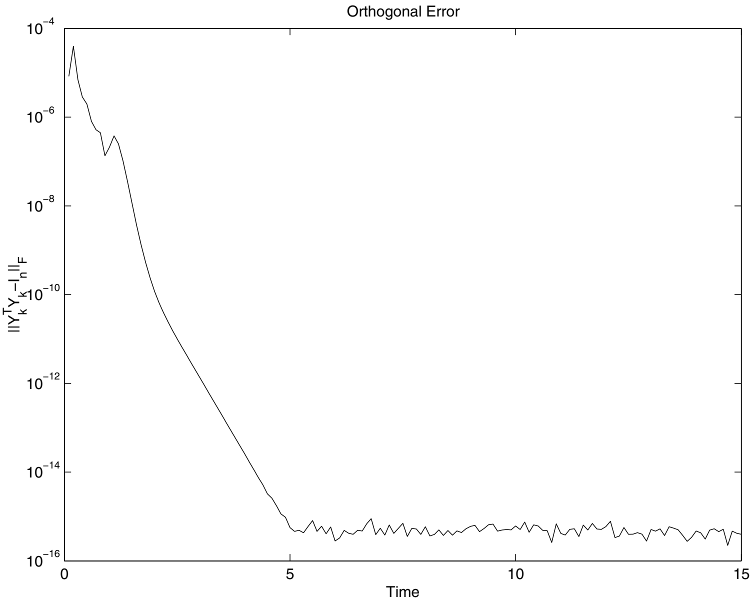

Fig. 1. Orthogonal error for Example 1.

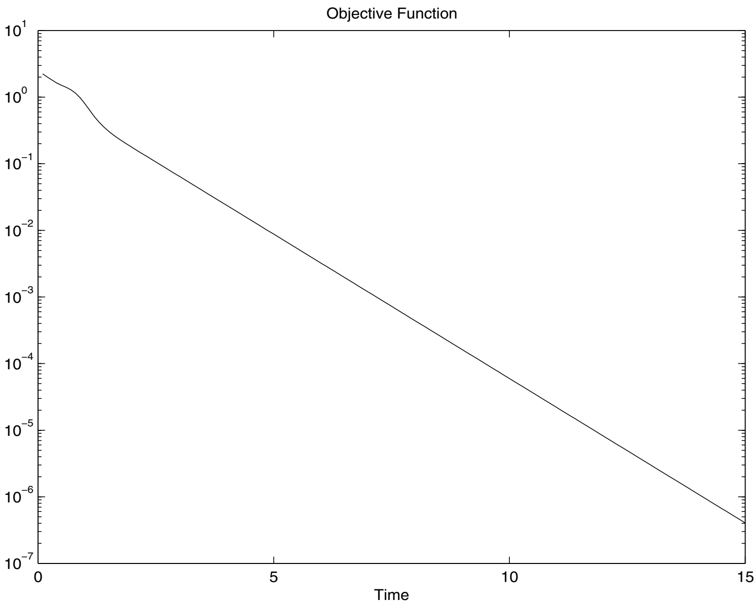

Fig. 2. Objective function for Example 1.

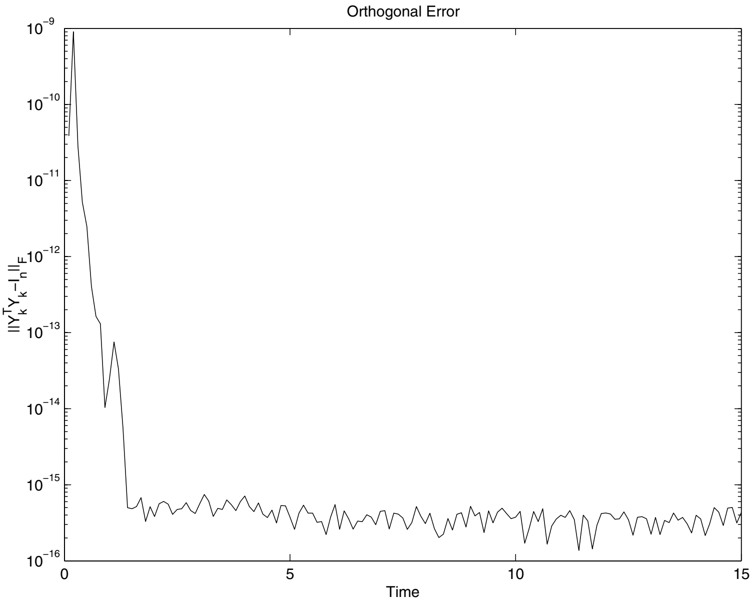

Fig. 3. Orthogonal error for Example 1.

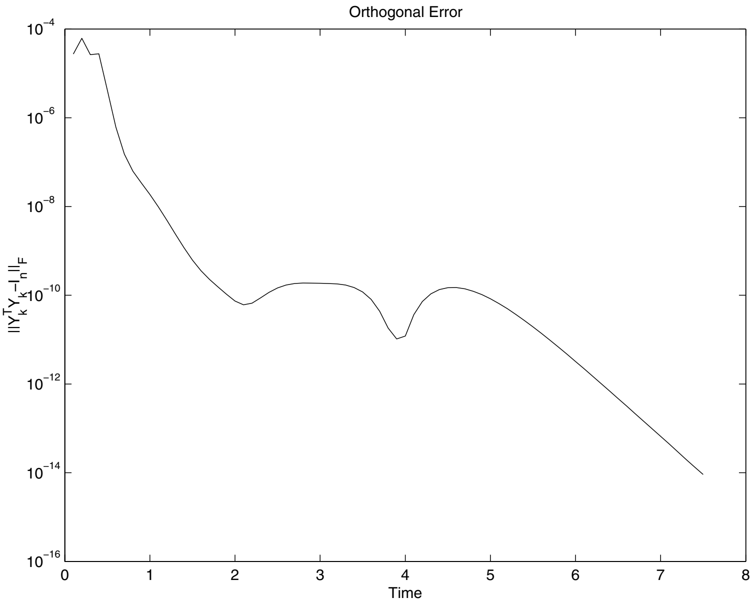

Fig. 4. Orthogonal errors for Example 2.

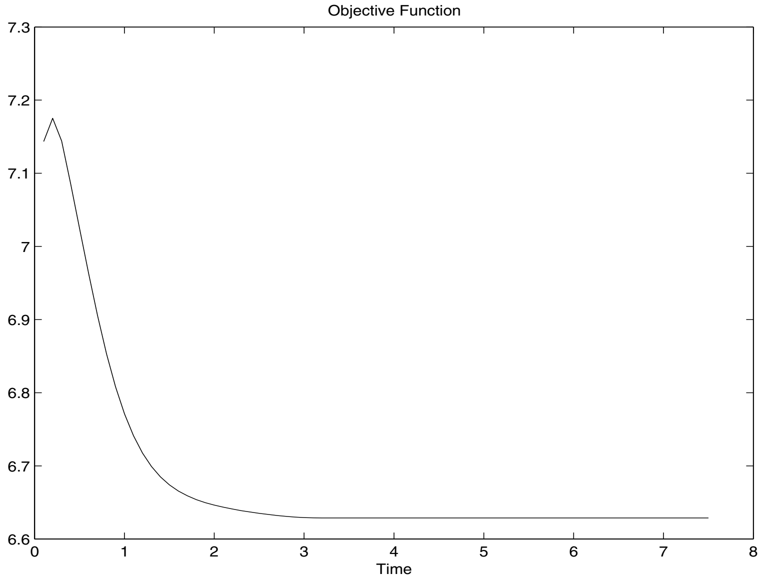

Fig. 5. Objective function for Example 2.

References

1. Ascher U.M.; Stabilization of invariants of discretized differential systems. Numer. Algorithms **14** (1997) 1–24
2. Ascher U.M., Chin H., Reich S.: Stabilization of DAEs and invariant manifolds. Numer. Math. **14** (1994) 131–149
3. Baumgarte J.: Stabilization of constraints and integrals of motion in dynamical systems. Comp. Math. Appl. Mech. Eng. **1** (1972) 1–16
4. Bridges T.J., Reich S.: Computing Lyapunov exponents on a Stiefel manifold. Physica D **156** (2001) 219–238
5. Chu M.T., Driessel K.R.: The projected gradient method for least squares matrix approximations with spectral constraints. SIAM J. Numer. Anal. **47** (1990) 1050–1060
6. Diele F., Lopez L., Peluso R.: The Cayley transform in the numerical solution of unitary differential systems. Adv. in Comp. Math. **8** (1998) 317–334
7. Higham D.J.: Time-stepping and preserving orthonormality. BIT **37** (1997) 24–36
8. Lopez L., Politi T.: Applications of the Cayley approach in the numerical solution of matrix differential systems on quadratic groups. Appl. Num. Math. **36** (2001) 35–55
9. Munthe-Kaas H.: Runge-Kutta methods on Lie groups. BIT **38** (1998) 92–111