Inverse Dynamics Controllers for Robust Control: Consequences for Neurocontrollers

Csaba Szepesvári[†] [‡] and András Lőrincz[†]

[†] Department of Adaptive Systems, Institute of Isotopes of the Hungarian Academy of Sciences, Budapest, P.O. Box 77, Hungary, H-1525
[‡] Bolyai Institute of Mathematics, University of Szeged, Szeged, Aradi vrt. tere 1, Hungary, H-6720

Abstract. It is proposed that controllers that approximate the inverse dynamics of the controlled plant can be used for on-line compensation of changes in the plant's dynamics. The idea is to use the very same controller in two modes at the same time: both for static and dynamic feedback. Implications for the learning of neurocontrollers are discussed. The proposed control mode relaxes the demand of precision and as a consequence, controllers that utilise direct associative learning by means of local function approximators may become more tractable in higher dimensional spaces.

1 Introduction

Neurocontrollers typically realise static state feedback control where the neural network is used to approximate the inverse dynamics of the controlled plant [1]. In practice it is often unknown *a priori* how precise such an approximation can be. On the other hand, it is well known that in this control mode even small approximation errors can lead to instability [2]. The same happens if one is given a precise model of the inverse dynamics, but the plant's dynamics changes. The simplest example of this kind is when a robot arm grasps an object that is heavy compared to the arm. This problem can be solved by increasing the stiffness of the robot, i.e., if one assumes a "strong" controller. Industrial controllers often meet this assumption, but recent interest has grown towards "light" controllers, such as robot arms with air muscles that can be considerably faster [3]. There are well-known ways of neutralising the effects of unmodeled dynamics, such as the σ-modification, signal normalisation, (relative) dead zone, and projection methods, being widely used and discussed in the control literature (see for example, [2]). Here we propose a new method where the feedforward controller is used in a parallel feedback operation mode thus resulting in robust control. The unattributed statements will be presented elsewhere.

2 Preliminaries

Let $\mathbf{R}^{m\times n}$ denote real $m \times n$ matrices. We say that a real matrix $\mathbf{A}$ admits a generalised inverse[1] if there is a matrix $\mathbf{X}$ for which $\mathbf{AXA} = \mathbf{A}$ holds [4]. For

[1] Sometimes it is called the pseudo- inverse, or simply the inverse of matrix $\mathbf{A}$.

convenience, the generalised inverse of a non-singular matrix $\mathbf{A}$ will be denoted by $\mathbf{A}^{-1}$. Assume that the plant's equation is given in the following form [5]:

$$\dot{\mathbf{q}} = \mathbf{b}(\mathbf{q}) + \mathbf{A}(\mathbf{q})\,\mathbf{u} \tag{1}$$

where $\mathbf{q} \in \mathbf{R}^n$ is the state vector of the plant, $\dot{\mathbf{q}}$ is the time derivative of $\mathbf{q}$, $\mathbf{u} \in \mathbf{R}^m$ is the control signal, $\mathbf{b}(\mathbf{q}) \in \mathbf{R}^n$, and $\mathbf{A}(\mathbf{q}) \in \mathbf{R}^{n\times m}$. We assume that the domain (denoted by D) of the state variable $\mathbf{q}$ is compact and is simply connected; that $n \leq m$, and for each $\mathbf{q} \in D$ the rank of matrix $\mathbf{A}(\mathbf{q})$ is equal to n; that is, the matrix is non-singular. As a consequence the plant is strongly controllable. In this case the inequality $n < m$ means that there are more independent actuators than state vector components, i.e., the control problem is redundant. Another kind of redundancy, or ill- posedness occurs when $n > m$ in which case even $\mathbf{A}^{-1}$ is non-unique. Further, we assume that both of the matrix fields, $\mathbf{A}(\mathbf{q})$ and $\mathbf{A}^{-1}(\mathbf{q})$ are differentiable w.r.t. $\mathbf{q}$.

Let $\mathbf{v} = \mathbf{v}(\mathbf{q})$ be a fixed n dimensional vector field over D. The *speed field tracking task* is to find the static state feedback control $\mathbf{u} = \mathbf{u}(\mathbf{q})$ that solves the equation

$$\mathbf{v}(\mathbf{q}) = \mathbf{b}(\mathbf{q}) + \mathbf{A}(\mathbf{q})\mathbf{u}(\mathbf{q}). \tag{2}$$

Conventional tasks, such as the *point to point control* and the *trajectory tracking* tasks cannot be exactly rewritten in the form of speed field tracking and speed field tracking is more robust against noise then these conventional tasks. Speed fields must be carefully designed if $\mathbf{A}(\mathbf{q})$ is singular (as in the case of a robot arm).

Given the plant's dynamics by Equation (1) the inverse dynamics of the plant may be written as follows:

$$\mathbf{p}(\mathbf{q},\dot{\mathbf{q}}) = \mathbf{A}^{-1}(\mathbf{q})\Big(\dot{\mathbf{q}} - \mathbf{b}(\mathbf{q})\Big) + \Big(\mathbf{E} - \mathbf{A}^{-1}(\mathbf{q})\mathbf{A}(\mathbf{q})\Big)\mathbf{y}(\mathbf{q},t), \tag{3}$$

where $\mathbf{E}$ is the unit matrix and $\mathbf{y} = \mathbf{y}(\mathbf{q},t)$ is an arbitrary function. Of course, the control signal $\mathbf{u}(\mathbf{q}) = \mathbf{p}(\mathbf{q},\mathbf{v}(\mathbf{q}))$ solves the speed field tracking control task. *In the following we will look at the main value of the inverse dynamics, i.e., we assume that* $\mathbf{y}(\mathbf{q},t) = 0$ *and thus we let* $\mathbf{p}(\mathbf{q},\mathbf{v}) = \mathbf{A}^{-1}(\mathbf{q})\Big(\dot{\mathbf{q}} - \mathbf{b}(\mathbf{q})\Big)$. This assumption simplifies the calculations and can be supported by appropriate learning algorithms.

3 Dynamic State Feedback control

The approximation errors of the inverse dynamics can be viewed as permanent perturbation to the plant's dynamics. Thus we assume that instead of Equation (1) the plant follows

$$\dot{\mathbf{q}} = \hat{\mathbf{b}}(\mathbf{q}) + \hat{\mathbf{A}}(\mathbf{q})\mathbf{u}, \tag{4}$$

where $\hat{\mathbf{A}}(\mathbf{q})$ is another nonsingular matrix field. Let us first assume that we seek a static state feedback compensatory control signal, $\mathbf{w} = \mathbf{w}(\mathbf{q})$, such that the control signal $\mathbf{u}(\mathbf{q}) + \mathbf{w}(\mathbf{q})$ solves the original speed field tracking problem for the perturbed plant.

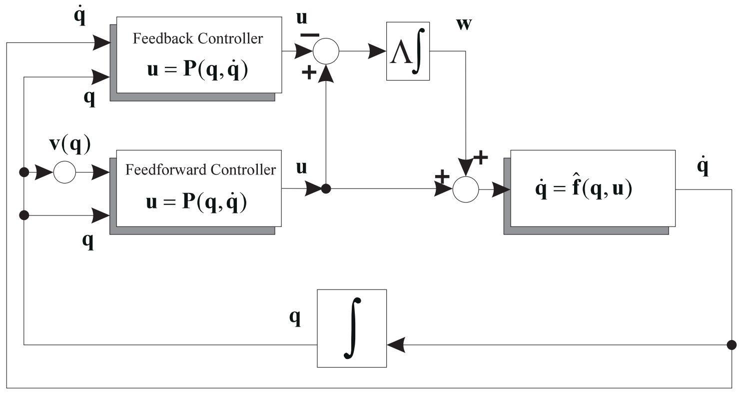

Fig. 1. SDS Control by doubling the role of the inverse dynamics controller

This can be solved by letting $\mathbf{w}(\mathbf{q})$ satisfy the equation $\mathbf{u}(\mathbf{q}) = \mathbf{p}\left(\mathbf{q}, \mathbf{v}\left(\mathbf{q}, \mathbf{w}(\mathbf{q})\right)\right)$, where $\mathbf{v}(\mathbf{q}, \mathbf{w}) = \dot{\mathbf{q}} = \hat{\mathbf{b}}(\mathbf{q}) + \hat{\mathbf{A}}(\mathbf{q})\Big(\mathbf{u}(\mathbf{q}) + \mathbf{w}\Big)$. Although $\mathbf{w}(\mathbf{q})$ can be explicitly expressed from this it contains terms like $\hat{\mathbf{A}}^{-1}(\mathbf{q})$ and $\hat{\mathbf{b}}(\mathbf{q})$ and thus to estimate $\mathbf{w}(\mathbf{q})$ on- line is approximately the same as retaining the adaptivity of the feedforward controller. This problem can be alleviated by introducing dynamic state feedback for estimating the compensatory control signal. The simplest corresponding error-feedback law is to let $\mathbf{w}$ change until $\mathbf{w}(\mathbf{q})$ satisfies the optimality condition:

$$\begin{aligned} \dot{\mathbf{w}} &= \Lambda\Big(\mathbf{u}(\mathbf{q}) - \mathbf{p}(\mathbf{q}, \mathbf{v}(\mathbf{q}, \mathbf{w}))\Big) \\ \dot{\mathbf{q}} &= \hat{\mathbf{b}}(\mathbf{q}) + \hat{\mathbf{A}}(\mathbf{q})\Big(\mathbf{u}(\mathbf{q}) + \mathbf{w}\Big), \end{aligned} \tag{5}$$

where Λ is a fixed positive number, the gain coefficient of dynamic feedback. If the speed of the plant is measurable then Equation System (5) can be realised by a compound control algorithm. The block diagram of the compound controller is given in Fig. 1. The compound controller will be called as the **S**tatic and **D**ynamic **S**tate Feedback Controller (SDS Controller).

If $\mathbf{S}$ is a symmetric real matrix then let $\lambda_{\min}(\mathbf{S})$ denote by the minimal eigenvalue of $\mathbf{S}$. Let $||\cdot||$ denote the Euclidean norm and let $\mathbf{z} = \hat{\mathbf{A}}(\mathbf{q})\mathbf{w} - \Big(\mathbf{v}(\mathbf{q}) - \hat{\mathbf{v}}(\mathbf{q})\Big)$. Simple calculation yields that $\dot{\mathbf{q}} = \mathbf{v}(\mathbf{q}) + \mathbf{z}$ and thus $\mathbf{z}$ is the error of tracking $\mathbf{v}(\mathbf{q})$. In order to guarantee the robustness of tracking it is sufficient to guarantee that the error of tracking is small. Conditions are given in the following theorem:

THEOREM 3.1 *Assume that the perturbation of* $\mathbf{A}(\mathbf{q})$ *can be decomposed as* $\hat{\mathbf{A}}(\mathbf{q}) = \mathbf{D}(\mathbf{q})\mathbf{A}(\mathbf{q})$. *Suppose that* $\mathbf{A}(\mathbf{q})$, $\mathbf{b}(\mathbf{q})$, $\mathbf{v}(\mathbf{q})$ *and* $\mathbf{D}(\mathbf{q})$, $\hat{\mathbf{b}}(\mathbf{q})$ *have continuous first* *derivatives and the constants* $a = \inf\{\, ||\mathbf{A}(\mathbf{q})|| \mid \mathbf{z} \in D\,\}$*,* $d = \inf\{\, ||\mathbf{D}(\mathbf{q})|| \mid \mathbf{z} \in D\,\}$*, and* $\lambda = \inf\{\, \lambda_{\min}(\mathbf{D}(\mathbf{q}) + \mathbf{D}^T(\mathbf{q})) \mid \mathbf{q} \in D\,\}$ *are positive. Then for all* $\epsilon > 0$ *there exists a gain* Λ *and an absorption time* $T > 0$ *such that for all* $\mathbf{z}(0)$ *that satisfy* $||\mathbf{z}(0)|| < K\Lambda$ *it holds that* $||\mathbf{z}(t)|| < \epsilon$ *provided that* $t > T$ *and the solution can be continued up to time* t*. Here* K *is a fixed positive constant and* $\mathbf{z}(0)$ *denotes the initial value of* $\mathbf{z}$*. Further,* $\Lambda \sim O(1/\epsilon)$*.*

For convenience $\mathbf{D}(\mathbf{q}) + \mathbf{D}^T(\mathbf{q})$ will be called the *Symmetrised Perturbation matrix* and will be abbreviated to *SP-matrix*. We say that a perturbation of Equation (1) is *non-invertive* or *uniformly positive definite* if λ is positive. The proof is based on a modification of Liapunov's Second Method with the semi-Liapunov function $V(\mathbf{x}) = \mathbf{z}^T\mathbf{z}$. If $\mathbf{z}$ satisfies the conclusions of Theorem 3.1 then it is said to admit the property of uniform ultimate boundedness (UUB).

Discussion

It is an important question whether the theory can be applied to real world situations. Along these lines it was shown that the SDS Control is stable when an idealised robot arm grasps or releases an idealised object (i.e., the mass of the end point changes). More recently, we have successfully applied this scheme to control a chaotic bioreactor with a severely mismatched inverse dynamics model. These results will be published elsewhere.

In neurocontrol the inverse dynamics is modelled by a neural network that may be inherently incapable of realising the inverse dynamics with zero error (i.e., the structural approximation error is non-zero). It is known that mere approximate direct inverse control does not guarantee the (ultimate) boundedness of the tracking error. Can then the boundedness of tracking error be guaranteed if one applies SDS Control? Assume that we approximate $\mathbf{A}^{-1}(\mathbf{q})$ and $\mathbf{b}(\mathbf{q})$ by $\mathbf{P}(\mathbf{q})$ and $\mathbf{s}(\mathbf{q})$, respectively. Then, of course, $\mathbf{A}(\mathbf{q})$ "approximates" $\mathbf{P}^{-1}(\mathbf{q})$. Now turning everything upside down let us imagine that the inverse dynamics of our controller is "exact", i.e., the plant's equation is given by $\dot{\mathbf{q}} = \mathbf{P}^{-1}(\mathbf{q}) + \mathbf{s}(\mathbf{q})$ and Equation (1) is thought of as the perturbed system. We can apply Theorem 3.1 and obtain stability provided that beside some smoothness conditions $\inf_{\mathbf{q}} \lambda_{\min}(\mathbf{D}^T(\mathbf{q}) + \mathbf{D}(\mathbf{q})) > 0$ holds, where $\mathbf{D}(\mathbf{q}) = \mathbf{A}(\mathbf{q})\mathbf{P}(\mathbf{q})$. If $\lambda = \inf_{\mathbf{q}} \lambda_{\min}(\mathbf{D}^T(\mathbf{q}) + \mathbf{D}(\mathbf{q})) > 0$, we say that *the controller represents the inverse dynamics of the plant signproperly.* Under this condition we have that for large enough gains the tracking error is UUB and the ultimate bound on error is proportional to $1/\Lambda$. This also means that *no matter how imprecise the controller is the error of tracking may be arbitrarily decreased by choosing large enough* Λ*s.* The positivity of the symmetrised perturbation matrix follows if $\mathbf{P}$ approximates $\mathbf{A}^{-1}$ sufficiently closely. Consequently, the number of parameters of the estimator can be reduced, which may revitalise the use of fast learning local approximator based neural networks as direct inverse controllers. Besides fast learning the advantage of local approximators is that *a priori* knowledge of the control manifold can be used in training, e.g. when using geometry representing networks one may introduce neighbour training without the loss of generality provided that the inverse dynamics function is smooth.

If the SP-matrix is uniformly positive definite then the use of SDS Control seems to be advantageous during the learning phase. However, if the initial controller does not represent the plant (semi-)sign properly then care is needed. There are two ways to achieve signproper initial estimates. First, one may initialise the controller so that it realises the everywhere zero function, or one may prelearn a 0th stage model until it becomes signproper. Then still one has to ensure that the learning algorithm preserves the signproper nature of approximation. For example, monotone learning admits this property.

When learning is *decoupled* from the control of plant (there is no reference signal during learning) then neither tracking, nor the tracking error and consequently nor SDS Control can be constructed. Thus in the following assume that learning is drived by a reference signal.

When *variational learning* is used (see, e.g. [6]), i.e., when the error of tracking is used for training the controller then SDS Control may delay or even may upset the identification of the inverse dynamics, since the inverse dynamics model may seem more precise than it is in reality. Another problem is that overcompensation may render the learning process unstable: large errors in trajectory tracking may be caused by both the overcompensation and the imprecise feedforward control signal. Consequently, this learning method should be cautiously used together with SDS Control. One solution may be to use temporal backpropagation that takes into account the double role played by the inverse dynamics of the controller by forcing weight sharing. Further research is needed to clarify this point.

In the case of *associative learning*, however, when the training data are given in the form of action-response pairs, it is possible to apply SDS Control during learning, but then the effective control signal should be used in to achieve proper learning. Note, that during learning the tracking of the desired trajectory is more precise with *SDS* Control than without it. This means that neither stability, nor the identification process is affected by SDS Control.

In the proofs we assumed that the perturbation is stationary, i.e., it does not change with time, which is unrealistic. It was shown however, that almost the same method can be applied to nonstationary perturbations. This means that if the changes are slow, i.e., these terms are bounded, then for large enough Λ one may retain the ultimate boundedness of the error signal. However, it is mentioned that in real world applications these changes are usually fast (e.g., when a robot arm grasps a heavy object). In such a case the error signal may become very large and the system may leave the stable region. This can be handled for example by the projection method.

Noise can be handled just as non-stationary perturbations provided that it has bounded amplitude and bandwidth. Note, that to the contrary to linear controllers the present method is capable of compensating such noise that enters the system at the inputs of the controller. However, if the noise enters the system just before the point where the compensatory vector is integrated, i.e., the noise affects $\dot{\mathbf{w}}$, then the system may easily become unstable. This problem, however, is not peculiar to our system, but is shared by every *dynamic* state feedback controller.

4 Conclusions

The so called SDS Control Mode was proposed to compensate inhomogeneous, non-linear, non-additive perturbations of non-linear plants that admit inverse dynamics. Such perturbations arise, for example, when a robot arm grasps or releases a heavy object. The SDS Controller is composed of two identical copies of an inverse dynamics controller. One copy acts as the original closed loop controller while the other identical copy is used to develop the compensatory signal. The advantage of this compound controller is that it can develop a control signal for compensating unseen perturbations and structural approximation errors and thus can control the plant more precisely than the closed loop feedforward controller alone. This relaxes the number of parameters required to achieve a given precision in control and thus may enable the use of fast learning local approximator based neural networks, that are otherwise known to suffer from combinatorial explosion in the dimension of the state space. SDS Control was implemented on a fully self- organising neural network controller that exploits only local, associative direct inverse estimation methods. Preliminary computer simulations show that compensation is fast for non-linear plants, which follows from the sketched theory for linear plants.

5 Acknowledgements

We are grateful to Prof. András Krámli for his invaluable comments and suggestions. This work was partially founded by OTKA grants T017110, T014330, T014566, and US-Hungarian Joint Fund Grant 168/91-A 519/95-A.

References

1. W.T. III Miller, R.S. Sutton, and P.J. Werbos, editors. *Neural Networks for Control.* MIT Press, Cambridge, Massachusetts, 1990.
2. R. Ortega and T. Yu. Theoretical results on robustness of direct adaptive controllers: A survey. In *Proc. 10th IFAC World Congress*, pages 26–31, Munich, 1987.
3. P. van der Smagt and K. Schulten. Control of pneumatic robot arm dynamics by a neural network. In *Proc. of the 1993 World Congress on Neural Networks*, volume 3, pages 180–183, Hillsdale, NJ., 1993. Lawrence Erlbaum Associates, Inc.
4. A. Ben-Israel and T.N.E. Greville. *Generalized Inverses: Theory and Applications.* Pure and Applied Mathematics, Wiley-Interscience. J. Wiley & Sons, New York, 1974.
5. A. Isidori. *Nonlinear Control Systems.* Springer-Verlag, Berlin, 1989.
6. K. Narendra and K. Parthasarathy. Gradient methods for the optimization of dynamical systems containing neural networks. *IEEE Trans. Neural Networks*, 2(2):252–262, 1991.