Principal Component Analysis in Accessible Control

Maximus McCune
maxmccune2024@u.northwestern.edu
Northwestern University, Shirley Ryan Ability Lab
Chicago, IL, USA

Andrew Thompson
athompson@u.northwestern.edu
Northwestern University, Shirley Ryan Ability Lab
Chicago, IL, USA

Brenna Argall
brenna.argall@northwestern.edu
Northwestern University, Shirley Ryan Ability Lab
Chicago, IL, USA

ABSTRACT

In a world where human-robot interactions are becoming increasingly common, it is important to ensure that our ability to interface with those robots is fully accessible to people with a wide variety of levels of ability. In noisy systems where the robot's motion is not state-dependent, Principal Component Analysis serves as a data-driven test to determine if a system is controllable, and to what degree it is. Its data-driven nature is applicable in cases where analytical models of the controller are very complex, such as neural networks or other machine learning models.

CCS CONCEPTS

• **Mathematics of computing** → *Probability and statistics*; • **Human-centered computing** → **HCI design and evaluation methods**; **Interaction design process and methods**.

KEYWORDS

Accessibility, Control Theory, Data Science, Statistics, Principal Component Analysis

ACM Reference Format:
Maximus McCune, Andrew Thompson, and Brenna Argall. 2024. Principal Component Analysis in Accessible Control. In *HRI '24: Assistive Applications, Accessibility, and Disability Ethics (A3DE), March 15, 2024, Boulder, CO.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

As robots become more present in our everyday world, one important question to consider is how they will interact with humans. Transparency in this interaction is essential in the understanding of how robots perform. One natural scale to consider these interactions is at the control interface level; and it is necessary to have interfaces that can function for people at a wide variety of ability levels[2][3]. People who have neuromotor dysfunctions often have difficulty operating traditional interfaces, such as video game controllers and haptic systems. Instead, their varying levels of residual motion present their own sets of challenges in designing control systems. To ensure that interfaces are accessible to people with neuromotor deficits, the control scheme for those interfaces must be designed in a manner which avails as much of the interface's control space as possible. By providing a user greater agency in their interface use, the systems which the user controls with that interface will be more transparent in function [1][7].

This paper was developed as a response to another study wherein participants with varying degrees of spinal cord injury were prompted to control a robotic arm via shoulder movements[4]. With each shoulder being restricted to simply performing up/down and forwards/backwards motions, the number of available degrees of freedom came into question. The four input degrees of freedom were theoretically not enough to span the 6-dimensional workspace of the robotic arm. However, this could not readily be verified because the mapping between the sensor readings and the robot behavior was determined via a variational auto-encoder network, and as a result, the complexity of the network and the high dimensionality of the data prevented simple controllability testing. Additionally, existing data-driven controllability tests only work for linear systems [6], which is a condition that does not hold for deep neural networks. The controllability test would therefore need to apply across both linear and nonlinear systems. We propose making use of Principal Component Analysis (PCA) to perform that testing. Specifically, we will analyze the controllability of systems designed for use by patients with limited residual movements, under what conditions the systems are and are not controllable, and how PCA can serve to answer those questions under a specific set of conditions.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 OBJECTIVE

The purpose of this paper is to propose an efficient, data-driven test for controllability in certain types of nonlinear systems with applications in the design of accessible control schema.

3 NOTATION

We will define the vector x to be the *state* of the system and the set X to be the set of all achievable states of the system. Expanding off of this, $\dot{x}$ will be the time derivative of the state, and $\dot{X}$ is the set of all possible time derivatives. Additionally, we will define $u(t)$ to be the vector control signal for the system and U to be the set of all possible control signals. Furthermore, we define the function $f : U \to \dot{X}$ to be our controller that maps the control inputs to the corresponding output behavior. All together, these create the form of the system that will be relevant to this paper, which is shown in Equation 1.

$$\dot{x} = f(u(t)) \tag{1}$$

Equation 1 will serve as the basis for all of the analysis conducted in Section 4, and will function with slight modification in Section 5.

4 NOISELESS CONTROLLABILITY TESTING

In performing traditional controllability tests for nonlinear systems, finding an analytical solution is known to be an incomputably difficult problem [8]. With the advent of neural networks and other highly nonlinear systems, it instead makes sense to take a more data-driven approach, where the model itself is not analyzed[6]. Rather, the input and output data will instead be analyzed, and then a result will be derived based off of that.

Next, we must define controllability. In this context, we will define it as when:

(A) That system starts at any arbitrary $x_0 \in X$.
(B) There exists a $u(t) \in U$ such that the final state of the system can become any arbitrary $x_f \in X$.

A more intuitive way to think about controllability is the system's ability to reach every point. For instance, if a robot arm could reach every possible end-effector orientation and position, then the system would be considered controllable. If, however, it could not reach a set of endpoints due to limitations in the controller, then that system would not be considered controllable.

By the Fundamental Theorem of Calculus, this is an equivalent statement to Equation 2, where t_0 and t_f are the initial and final time, respectively, and x_0 and x_f can be any arbitrary states.

$$x_f = x_0 + \int_{t_0}^{t_f} \dot{x} dt = x_0 + \int_{t_0}^{t_f} f(u(t))dt \tag{2}$$

Using this fact, we will prove in the case where the data in symmetric about the origin, a matrix rank test (MRT) is sufficient to demonstrate controllability. To do this, we will proceed with a proof by induction. Suppose that $u(t)$ is constant on some time interval $[t_0, t_1)$, such that $t_0 < t_1 < t_f$. It follows that Equation 2 can be rewritten as seen in Equation 3.

$$x_f = x_0 + \int_{t_0}^{t_f} f(u(t))dt = x_0 + (t_1 - t_0)f(u(t_0)) + \int_{t_1}^{t_f} f(u(t))dt \tag{3}$$

Defining the variable $x_1 = x_0 + (t_1 - t_0)f(u(t_0))$, Equation 3 further simplifies to the following as seen in Equation 4, where x_1 is trivially a linear combination of x_0 and $f(u(t_0))$.

$$x_f = x_0 + (t_1 - t_0)f(u(t_0)) + \int_{t_1}^{t_f} f(u(t))dt = x_1 + \int_{t_1}^{t_f} f(u(t))dt \tag{4}$$

Now, suppose $u(t)$ is piecewise constant in n different regions throughout the entire interval $[t_0, t_f]$. By the result derived above, it follows that Equation 4 would be equivalent to Equation 5.

$$x_f = x_0 + (t_1 - t_0)f(u(t_0)) + \int_{t_1}^{t_f} f(u(t))dt = x_0 + \sum_{i=0}^{n}(t_{i+1} - t_i)f(u(t_i)) \tag{5}$$

Since $u(t)$ is assumed to be continuous, it then follows that in the limit as n approaches infinity Equation 5 would hold for any arbitrary signal $u(t)$. Therefore, in the limit $x_f - x_0$ is equivalent to a certain linear combination of elements of $\dot{X}$. As we are dealing with linear combinations, this structure nicely lends itself to a linear algebra-based data-driven test, and in our case, the aforementioned MRT.

Since $x_f - x_0$ is a linear combination of values of f, it follows that if the values of f serve as a basis for the space X, the system is controllable. To calculate this, all one needs to do is compute the rank of the matrix $[f(u(t_0)), f(u(t_1)), ...f(u(t_f))]$. If it is equal to the dimension of X, then the values of f are a valid basis, and the system is therefore controllable.

One important caveat here is that the coefficients for the linear combinations are all positive, as it is not possible to hold a value of $u(t)$ for negative time. Therefore, if the data in the workspace is not centered about zero, the result of this test may be erroneous.

Another important limitation is that through this test alone, it is not possible to identify axes with problematic data. Therefore it is not possible to identify which steps need to be taken to induce desirable behavior. Principal Component Analysis (PCA) addresses this issue by generalizing the idea of matrix rank[5]. It creates a series of axes that span a dataset and computes the variance across each axis. To perform an MRT from PCA, all one would need to do is count the number of non-zero variances, meaning that PCA does provides all of the information that the MRT does. However, PCA can also address another issue which has gone undiscussed: noise.

5 NOISY CONTROLLABILITY TESTING

In dealing with real-world systems, noise is an ever-present issue. To model it in this context, we will add $\epsilon\eta(t)$ to the control signal $u(t)$, where $\eta(t)$ is some arbitrary perturbation and ϵ is some small number. Therefore, the system shown in Equation 1 becomes the more complex Equation 6.

$$\dot{x} = f(u(t) + \epsilon\eta(t)) \tag{6}$$

Due to the lack of causal connection between η and u, we cannot make use of the matrix rank test seen in Section 4. The introduction of η will instead induce a small perturbation across every axis, making virtually every measurable system have rank of sufficient order, irrelevant of its true controllability. Instead, we will make use of Principal Component Analysis (PCA) to determine exactly which axes have a significant amount of variance on them, and which axes have variance that is simply due to the noise from η.

To determine which axes' variance is noise-induced, we propose making use of a 'noise threshold,' which we decided to be around 1% of the variance across the largest principal axis. If the variance across any axis was found to be roughly less than or equal to this value, they should be treated as noise and dealt with accordingly. This offers the advantage of not only finding noise in the data, but also finding exactly which axes are affected.

Another advantage offered by PCA is its ability to determine the distributions of the data across each axis, which was another limitation of the MRT. The data can be directly projected onto the principal axes that PCA constructs, and the distribution along each axis can be easily determined. This can show if the system is symmetric, and therefore if it is controllable.

6 RESULTS

To determine the efficacy of the above tests, systems that were both controllable and uncontrollable were simulated with a variety of input data. One such controllable noiseless system is show below in Equation 7, with the control input shown in Equation 8:

$$f(\vec{u}) = \vec{u} \tag{7}$$

$$u(t) = \begin{bmatrix} 2sin(t) \\ \frac{t}{2} \end{bmatrix} \tag{8}$$

In the noisy equivalent of Equation 7, a small $\epsilon\eta$ was added to the output of f. The results of the MRT and PCA controllability tests are shown below in Table 1.

Test	Noiseless	Noisy
MRT	Controllable	Controllable
PCA	Controllable	Controllable

Table 1: The results of the controllability test for the noisy and noiseless controllable sytems

System	PC 1	Variance 1	PC 2	Variance 2
Noiseless	$\begin{bmatrix} -.349 \\ -.947 \end{bmatrix}$	2.140	$\begin{bmatrix} -.937 \\ .349 \end{bmatrix}$	1.724
Noisy	$\begin{bmatrix} -.349 \\ -.947 \end{bmatrix}$	2.139	$\begin{bmatrix} -.937 \\ .349 \end{bmatrix}$	1.724

Table 2: The results of PCA for the noisy and noiseless nonlinaer controllable system

An uncontrollable system that was tested is shown below in Equation 9, with the control input shown in Equation 8.

$$f(\vec{u}) = \begin{bmatrix} u_1 \\ 0 \end{bmatrix} \tag{9}$$

In the noisy equivalent of Equation 9, a small $\epsilon\eta$ was added to the output of f. The results of the MRT and PCA controllability tests are shown below in Table 3.

Test	Noiseless	Noisy
MRT	Uncontrollable	Controllable
PCA	Uncontrollable	Uncontrollable

Table 3: The results of the controllability test for the noisy and noiseless uncontrollable sytems

Notably, both of the above systems discussed are linear. PCA also can identify controllability in nonlinear systems, such as the controllable nonlinear system seen below in Equations 10 and 11.

$$\vec{u}(t) = \begin{bmatrix} t \\ cos(t+3) \end{bmatrix} \tag{10}$$

$$f(\vec{u}) = \begin{bmatrix} sin(u_1) \\ u_2 \end{bmatrix} \tag{11}$$

System	PC 1	Variance 1	PC 2	Variance 2
Noiseless	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	1.856	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	0
Noisy	$\begin{bmatrix} .999999 \\ .00003 \end{bmatrix}$	1.775	$\begin{bmatrix} -.00003 \\ .99999 \end{bmatrix}$	0.00008

Table 4: The results of PCA for the noisy and noiseless linear uncontrollable system

Once again, in the noisy equivalent of the system from Equations 10 and 11, we add a small $\epsilon\eta$ to the control signal u. The results of the test are shown below in Table 5, with more data from the PCA test being shown in Table 6.

Test	Noiseless	Noisy
MRT	Controllable	Controllable
PCA	Controllable	Controllable

Table 5: The results of the controllability test for the noisy and noiseless controllable nonlinear system

System	PC 1	Variance 1	PC 2	Variance 2
Noiseless	$\begin{bmatrix} -.535 \\ -.844 \end{bmatrix}$	.581	$\begin{bmatrix} -.844 \\ .535 \end{bmatrix}$	.388
Noisy	$\begin{bmatrix} -.540 \\ -.841 \end{bmatrix}$	.584	$\begin{bmatrix} -.841 \\ .540 \end{bmatrix}$	.386

Table 6: The results of PCA for the noisy and noiseless nonlinaer controllable system

These tests can also identify uncontrollability in nonlinear systems. Using the same control input found in Equation 10, we can modify the system found in Equation 11 to an uncontrollable variant, found in Equation 12.

$$f(\vec{u}) = \begin{bmatrix} sin(u_1) \\ 2sin(u_1) \end{bmatrix} \tag{12}$$

This system provides a controller whose mapping from the control signal to the control signal is nonlinear, but the two outputs are linear combinations of each other. This leads to uncontrollability, and the results of the tests performed on the noisy and noiseless systems are shown in Table 7.

Test	Noiseless	Noisy
MRT	Uncontrollable	Controllable
PCA	Uncontrollable	Uncontrollable

Table 7: The results of the controllability test for the noisy and noiseless uncontrollable nonlinear system

System	PC 1	Variance 1	PC 2	Variance 2
Noiseless	$\begin{bmatrix} .447 \\ -.894 \end{bmatrix}$	2.218	$\begin{bmatrix} -.894 \\ -.447 \end{bmatrix}$	0
Noisy	$\begin{bmatrix} .447 \\ -.894 \end{bmatrix}$	2.218	$\begin{bmatrix} -.894 \\ -.447 \end{bmatrix}$	0.01

Table 8: The results of PCA for the noisy and noiseless nonlinear uncontrollable system

7 DISCUSSION

Through the analyses of these systems, we found that the MRT was successful in correctly identifying some of the noiseless systems as controllable or uncontrollable, only failing when the data was not symmetric about the origin. However, it identified every noisy option as controllable regardless of the system's true nature. In contrast, PCA was found to work in all of the observed cases. It could successfully identify the axes that were controllable and uncontrollable, any asymmetric axes in the dataset, and even functioned when the data had artificially induced noise. Additionally, both tests performed as expected, irrespective of the linearity of the system.

The authors found that no standardized threshold of variance attributable to noise (as discussed in Section 5) could be found. Therefore, the authors of this paper decided to select the value of 1% of the maximum variance as the threshold to be attributed to noise, and anything more than that could be considered statistically significant.

One limitation of the PCA test is it requires prior knowledge of the workspace and requires the workspace to be euclidean. Additionally, it does only apply to systems in the form of Equation 1, meaning it cannot easily be generalized to a generic nonlinear case. However it does provide unique insights into the dynamics of the system, which is something that is generally difficult to achieve when dealing with deep neural networks.

8 CONCLUSION

When analyzing control systems designed for people with different levels of residual motion, accessibility is important to consider. If the system is of the form seen in Equation 6, we propose using PCA on the output dataset to determine if the system is controllable. PCA provides insights on which axes are not controllable, to what degree they are not controllable, and possibly even steps to correct for those errors.

REFERENCES

[1] Victoria Alonso and Paloma De La Puente. System transparency in shared autonomy: A mini review. *Frontiers in Neurorobotics*, 12(November):1–11, 2018.
[2] Mahdieh Nejati Javaremi, Michael Young, and Brenna D. Argall. Interface operation and implications for shared-control assistive robots, 2019.
[3] Raida Karim, Amal Nanavati, Taylor A Kessler Faulkner, and Siddhartha S Srinivasa. Investigating the Levels of Autonomy for Personalization in Assistive Robotics. 2023.
[4] Jongmin M. Lee, Temesgen Gebrekristos, Dalia De Santis, Mahdieh Nejati-Javaremi, Deepak Gopinath, Biraj Parikh, Ferdinando A. Mussa-Ivaldi, and Brenna Argall. An exploratory multi-session study of learning high-dimensional body-machine interfacing for assistive robot control. In *Proceedings of the IEEE-RAS-EMBS International Conference on Rehabilitation Robotics (ICORR)*, 2023.
[5] Andrzej Maćkiewicz and Waldemar Ratajczak. Principal components analysis (pca). *Computers Geosciences*, 19(3):303–342, 1993.
[6] Vikas Kumar Mishra, Ivan Markovsky, and Ben Grossmann. Data-driven tests for controllability. *IEEE Control Systems Letters*, 5(2):517–522, 2021.
[7] Jef Raskin. Intuitive equals familiar, 1994.
[8] Henk J. van Waarde, Jaap Eising, Harry L. Trentelman, and M. Kanat Camlibel. Data informativity: A new perspective on data-driven analysis and control. *IEEE Transactions on Automatic Control*, 65(11):4753–4768, 2020.

Received 16 February 2024