MODEL-ORDER SELECTION FOR THE FLUCTUATION ANALYSIS
OF HUMAN GAIT DATA

A Thesis in
Engineering Science and Mechanics

by
Ruolin Shi

© 2016 Ruolin Shi

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

December 2016
The thesis of Ruolin Shi was reviewed and approved* by the following:

Joseph P. Cusumano
Professor of Engineering Science and Mechanics
Thesis Advisor

Corina S. Drapaca
Associate Professor of Engineering Science and Mechanics

Judith Todd Copley
Department Head of Engineering Science and Mechanics

*Signatures are on file in the Graduate School
ABSTRACT

Variability in the repeated execution of human movements is generally believed to contain key information for understanding the neuromotor system. A data analysis framework based on the idea of goal equivalent manifold (GEM) has been previously developed to study inter-trial error correction and to predict both the temporal and geometric structure of variability for skilled motor performance. The GEM framework successfully provides a general model-based approach for characterizing the neuromotor dynamic system from experimental data. However, the first-order autoregressive (AR) models adopted by previous studies have never been shown to be adequate. Furthermore, they are incapable of explaining why there is more variability in the performance of older people, especially given the finding that both young and old people share a similar geometrical structure to their variability. In this thesis, we focus on human gait data introducing higher-order AR model into the GEM framework so that we can look more carefully at the inter-stride dynamics, particularly in regard to the required state space dimension needed to model fluctuations. We use the Akaike and Bayesian information criteria (AIC & BIC) to analyze the distribution of preferred model orders of two age groups. We find that the most probable model orders of two groups are no larger than three, especially, the model orders selected by BIC indicate the first-order AR model could describe the dynamic system well by showing that first and second order models are most frequently picked. For geometric structure, a lower order model is preferred for variability in goal-relevant subspace while a higher order model is preferred in null subspace. By performing permutation test, results of statistical difference between preferred model orders of two age groups do not show consistency between models selected by AIC and BIC, for which we cannot make a firm conclusion and further studies may be required.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter 1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 GEM framework</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Time Series Analysis</td>
<td>4</td>
</tr>
<tr>
<td>Chapter 2 Data Analysis Methodology</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Treadmill Gait and GEM</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Modeling Inter-Stride Fluctuations</td>
<td>11</td>
</tr>
<tr>
<td>2.3 Building AR Model and Model Selection</td>
<td>13</td>
</tr>
<tr>
<td>2.4 Significance Test</td>
<td>17</td>
</tr>
<tr>
<td>2.5 Data Processing in Software Package</td>
<td>18</td>
</tr>
<tr>
<td>2.6 Subjects and Experiment Setup</td>
<td>19</td>
</tr>
<tr>
<td>Chapter 3 Results</td>
<td>20</td>
</tr>
<tr>
<td>3.1 Variability of Young and Old Groups</td>
<td>20</td>
</tr>
<tr>
<td>3.2 Results regarding overall group</td>
<td>21</td>
</tr>
<tr>
<td>3.3 Results regarding Normal and Tangential Directions</td>
<td>22</td>
</tr>
<tr>
<td>3.4 Results regarding Dynamical Geometry respected to GEM</td>
<td>25</td>
</tr>
<tr>
<td>Chapter 4 Conclusion and Discussion</td>
<td>27</td>
</tr>
<tr>
<td>References</td>
<td>30</td>
</tr>
<tr>
<td>Appendix</td>
<td>33</td>
</tr>
</tbody>
</table>

- Module for calculating Body-Goal variability matrix | 33 |
- Module for finding AIC and BIC for each TrD | 34 |
LIST OF FIGURES

Figure 2-1 Goal Equivalent Manifold (GEM) for treadmill walking task ............................. 10
Figure 2-2 A typical MIP controller with weak POP term in body state space ..................... 13
Figure 3-1 Variability of Young Participants and Old Participants .................................. 21
Figure 3-2 Treadmill walking model order selection and significance test .......................... 22
Figure 3-3 Treadmill walking model order selection and significance test in goal relevant subspace ....................................................................................................................... 23
Figure 3-4 Treadmill walking model order selection and significance test in null subspace .. 24
Figure 3-5 Dynamical geometry structure with respect to GEM of young group ............... 25
Figure 3-6 Dynamical geometry structure with respect to GEM of old group ................. 26
ACKNOWLEDGEMENTS

First and foremost, I wish to express my sincere gratitude to my advisor, Prof. Joseph P. Cusumano, professor of department of Engineering Science and Mechanics. His patience, motivation, and immense knowledge have supported me academically and emotionally through the rough road to finish this thesis. He helped me come up with the thesis topic and guided me through the model development. I could not have imagined having a better advisor and mentor for my study.

Besides my advisor, I would like to thank Dr. Corina S. Drapaca for being the member of my thesis committee and providing insightful comments and challenging questions. Also, thank you to Judith Todd Copley, Department Head of Engineering Science and Mechanics, for providing all the supports and conveniences I need to finish my thesis.

I would also like to thank Jonathan B. Dingwell and his group for sharing the experiment data of treadmill walking used in this thesis.

Last but not the least, I would like to thank my family: my parents Jianwei Shi and Changhong Yang. Their financial and spiritual supports have been unconditional throughout my life, especially for the period I studied in the Pennsylvania State University.
Chapter 1

Introduction

Movement variability has been used clinically to diagnose the health of the nervous system for a long time. It is considered to be a key characteristic of the biological dynamic system underlying human goal-directed movement. One conceptual framework grounded on the idea of goal equivalent manifold (GEM) has been previously established to study inter-trial error correction, and to predict both the temporal and geometric structure of variability for skilled motor performance (Cusumano & Cesari, 2006). The GEM framework successfully provides a general model-based approach for characterizing the neuromotor dynamic system from experimental data.

However, in the study of gait of treadmill walking, the first-order autoregressive (AR) models adopted by previous studies (Dingwell, John, & Cusumano, 2010) are incapable of explaining why there is more variability in the performance of older people, especially given the finding that both young and old people share a similar geometrical structure to their variability.

Motivated by addressing above anomaly, this thesis applied two general ideas to study the variability in gait analysis. The first one is GEM framework, and the second idea is time series analysis. Experimental data of skilled performance near a task manifold can be treated as a time series. Finding the best fit model for the time series provides insight of variability geometry. We introduce higher-order AR models into GEM framework so that we can look more carefully at the inter-stride dynamics, particularly in regard to the required state space dimension. Akaike and Bayesian information criteria (AIC & BIC) (Akaike, 1973; Schwarz, 1978) are used to find the best fit model, and permutation test (Collingridge, 2013) is carried out to show significance of our findings.
We find that the most probable best fit model orders selected by AIC are typically larger than the one selected by BIC. However, both criteria show the most probable order of two groups are no larger than three, especially the model orders selected by BIC show that first and second order models are most frequently picked. For geometric structure, a lower order model is preferred for variability in goal-relevant subspace while a higher order model is preferred in null subspace. More specifically, the BIC picked preferred model order of goal-relevant subspace equals to one, which indicate the first-order AR model in previous study is capable to describe and provide insights of the dynamic system.

When we address the difference between young and old groups, results of statistical difference between preferred model orders do not show consistency between models selected by AIC and BIC, for which we cannot firmly claim whether young subjects tend to have lower model orders or not. We state our findings and discuss the probable explanations in later chapters. Further studies may be required.

Here, we review the main concepts in GEM framework and time series analysis.

1.1 GEM framework

The GEM framework is previously developed to unify several concepts of redundant system control which characterizing error correcting processes for skilled task performance. The variability observed for many movement tasks, like reaching to a target and walking, are explained well with GEM framework.

For a typical goal-oriented task, the motor system in humans is redundant, which means there are more than one combination of body state (e.g. joint speed, joint angle, etc.) can fully complete the task (Bernstein, 1967). Task manifolds can be used to study the redundancy in
neuromotor system. Task manifolds are surfaces in a body state space that contain all the possible states that can perfectly execute the task.

To experimentally apply the idea of task manifolds, there are two main approaches. One is uncontrolled manifold (UCM) analysis. It defines the task manifold using the trajectory of a given movement at each instant. The task’s goal is typically not used to define the task manifold. Instead, UCM analysis takes the average movement in a set of trials to represent the goal. By definition of a task manifold, variations along the manifold, in other words, that are tangential to it, will not affect the execution of the task (Cusumano & Dingwell, 2013). Conversely, variations normal to the manifold will lead to performance error (Latash & Scholz, 2002). Thus, the UCM approach is based on the hypothesis that only deviation in the normal direction will be regulated, and leaving the tangential deviations uncontrolled.

Another approach based on task manifolds tolerance-noise-covariation (TNC) analysis. It decomposes observed performance body-level variability into three goal-level costs which related to three different solution manifolds (Cohen & Sternad, 2009; Ranganathan & Newell, 2010). The tolerance cost relates the goal-level error to body-level variability by measuring the amplification ratio between them. The noise cost measures the amplitude of overall body-level fluctuation on goal-level error. The covariation cost focuses on the orientation and alignment of variability with respect to the task manifold. Different from UCM, the TNC approach represents task manifolds is a minimal space of variables required to specify task execution. The TNC approach has been used to explain how human learn to reach optimal performance by minimizing these three costs.

One focus of movement control in nervous system is to understand its optimization processes (Engelbrecht, 2001; Flash & Hogan, 1985; Hoyt & Taylor, 1981; Zarrugh, Todd, & Ralston, 1974). Average behavior, instead of variability, is explained by most existing optimization approaches. However, they never clarify whether the variability is regulated as impediment to
motor performance (Harris & Wolpert, 1998; Körding & Wolpert, 2004; Scheidt, Dingwell, & Mussa-Ivaldi, 2001), or as a target to maximize task performance (Todorov & Jordan, 2002).

To address above issues, stochastic optimal control theory has been proposed as the basis of computational motor control models which can be both explanatory and predictive. In addition, by combining task geometry ideas and stochastic optimal control theory, the minimum intervention principle (MIP) provides computational framework to predict regulation behavior in redundant motor systems (Todorov, 2004).

Local stability of a system is defined as its response to sufficiently small perturbations to a steady state. One system, which can persist small perturbations near its steady state, is stable, otherwise it is unstable (Verhulst, 1996). In addition, local stability study can also quantify the effectiveness of a controller. Measuring variability alone does not provide direct information of local stability since perturbations cannot quantify the response of dynamic system.

1.2 Time Series Analysis

A time series is a sequence of data points listed in order of time. Analyzing a time series is to find the internal structure (such as autocorrelation, trend or seasonal variation) that data points taken over time may have.

Generally, a stochastic model for time series analysis will use the natural ordering of time so that the information in the past can be used to derive the values in the future, rather than the other direction. In addition, the fact that values in a small range of time, tend to have a stronger relationship between values further apart.

We used time-domain parametric methods in this thesis to extract dynamical information from gait data. The assumption of parametric methods is that the structure of the underlying process can be described by a set of parameters.
A well-known parametric method, the Box-Jenkins method, applies autoregressive moving average (ARMA or ARIMA) models to find the best structures and parameters that can fit the past values of a time series (Box & Jenkins, 1970). In this thesis, vector autoregressive (VAR) models are chosen to perform time series analysis because of the redundancy in motor system of humans. Instead of taking scalar variables in Box-Jenkins methods, VAR models generalize the univariate autoregressive model (AR model) by taking vectors as evolving variable (Hacker & Hatemi-J, 2008). A \( p \)-th order has the form of

\[
\mathbf{x}_k = \mathbf{c} + \mathbf{B}_1 \mathbf{x}_{k-1} + \mathbf{B}_2 \mathbf{x}_{k-2} + \cdots + \mathbf{B}_p \mathbf{x}_{k-p} + \mathbf{e}_k
\]  \hspace{1cm} (1-1)

where the \( i \)-period back observation \( \mathbf{x}_{k-i} \) is called the \( i \)-th lag of \( \mathbf{x} \), \( \mathbf{c} \) is a vector of constants, \( \mathbf{B}_i \) is a time-invariant matrix and \( \mathbf{e}_k \) is a vector of error terms. A VAR model which has the form of Eq. 1-1 is called a \( p \)-th order VAR or a VAR with \( p \) lags.

While first-order VAR model used in previous GEM study fits and explains experimental data well, there are still some questions remaining unanswered. The main criticism is regarding whether the first-order VAR model is complex enough to describe the underlying processes. We replace the first-order VAR model with a set of VAR models with different lags to promote GEM framework the ability to model the process with higher dimensions. We quote VAR as AR, except for notice in later chapter since all data in our analysis is vector.

In process of analyzing time series, usually a set of models with different parameter settings (in our case, the parameters are lag order and coefficient matrix) will be constructed. A criteria is needed here to measure the fit of each setting so that we are able to select the appropriate model to describe the underlying stochastic process. In our study, two regular information criteria have been chosen. One is the Akaike information criterion (AIC) (Akaike, 1973) and the other one is the Bayesian information criterion (BIC) (Schwarz, 1978).

\[
AIC = 2j - 2 \ln(\hat{L})
\]  \hspace{1cm} (1-2)
\[ BIC = -2\ln(\hat{L}) + j\ln(n) \]  

(1-3)

where \( \hat{L} \) is the value of the likelihood of certain parameter sets (Edwards, 1972), it equals to the probability of the observed outcomes given the value of parameters. \( n \) denotes the sample size and \( j \) denotes the number of parameters.

The Akaike information criterion (AIC) is a measure of the goodness of fit of an estimated statistical model. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a means for model selection.

AIC is based on information theory: it estimates the relative information loss generated during the data generation process using a given model. Usually a more complex model would fit data better (Myung, 2000), however, the goodness of fit sometimes may not worth the trade-off with the added model complexity. One notable property of AIC is that AIC does not provide information about the absolute quality of the model. In the case of no well-fitted model exists, there is no sign to warn that.

The Bayesian information criterion (BIC) or Schwarz criterion is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. It is based, in part, on the likelihood function and it is closely related to the Akaike information criterion (AIC).

When fitting models, it is possible to increase the likelihood by adding parameters, but doing so may result in overfitting (Richman, Staszewski, & Simon, 1995). Both BIC and AIC resolve this problem by introducing a penalty term for the number of parameters in the model (Schunn & Wallach, 2005); the penalty term in BIC is larger than it in AIC.

AIC and BIC were originally derived under different assumptions and are useful in different settings. AIC was derived under the assumption that the true model requires an infinite number of parameters and attempts to minimize the information loss by using a given finite-dimensional model to approximate it. BIC was derived as a large-sample approximation to Bayesian selection among a fixed set of finite dimensional models. This difference leads to different
properties. AIC always has a chance of choosing too big a model, regardless of sample size $n$. BIC has very little chance of choosing too big a model if $n$ is sufficient, but it has a larger chance than AIC, for any given $n$, of choosing too small a model (Acquah, 2010).

In the following chapters, we first demonstrate the theoretical methodology of GEM framework and model selection. Then, we briefly describe the conditions of how the treadmill walking experiment is conducted, and how we process data using software package. Next, we will show and compare the order distribution of best-fit model of both young and old groups.
Chapter 2

Data Analysis Methodology

In the first two sections of this chapter, we review the processes to set up a goal equivalent task manifold and model inter-stride fluctuations. In the next two sections, we demonstrate the establishment of the multi-order autoregressive model for treadmill walking task, the related model selection process, and the significance test method for our study. Two hypothesis are made trying to explain variability difference between two age groups. In the last two sections, we clarify the software platform we used and the experiment condition which our data have been collected.

2.1 Treadmill Gait and GEM

To relate the goal-level performance to a body state for certain task, first thing to do is to build a goal function.

\[ f(x; p) = 0 \]  

where \( f \) is a vector goal function with dimension \( D_e \), \( x \) here is a body state vector with dimension \( D_b \). Vector \( p \) specifies the environment parameters (e.g., the speed of a treadmill in the walking task). By definition, all the body state vectors \( x \) which satisfy Eq. 2-1 are able to reach the goal equally. Thus, \( x \) is given the name execution variables.

For example, in treadmill walking task, the top requirement for our tester is to stay on the treadmill, so that we can study the fluctuations about the steady state. Each step contributes to the position change differently, while the summation of all the changes should be no more than the half length of a treadmill. Assuming the belt is working at speed \( v \), the net change for each step can be
represented as $L_n - vT_n$. Where $L_n$ is the stride length and $T_n$ is the stride time. The task goal could be rephrased in a mathematical way by:

$$\left| \sum_{n=1}^{N} (L_n - vT_n) \right| < \frac{L_{\text{Treadmill}}}{2}$$

(2-2)

Different strategies can be used to generate qualified $[L_n T_n]$ pairs. One of the simplest strategy is to assume the net change of each step maintains zero:

$$L_n - vT_n = 0 \rightarrow \frac{L_n}{T_n} = v$$

(2-3)

Here we can easily find the goal function becomes

$$f(L, T; v) = \frac{L}{T} - v = 0$$

(2-4)

From Eq. 2-4, we have $x = [L, T]^T$ and $p = [v]$. Note that $D_x=1$ and $D_b=2$, typically $D_x < D_b$ since the redundancy exists in the human body. All the combination of $L$ and $T$ which satisfy Eq. 2-3 will keep the net change of each step to zero. We refer the set of execution variable $x$ like that to a goal equivalent manifold $\mathcal{G}$:

$$\mathcal{G} = \{ x | f(x; p) = 0 \}$$

(2-5)

in our case, $\mathcal{G}$ is a straight line as shown in Figure 2-1.

So far, we generated a goal equivalent manifold. The whole process is more like a TNC approach rather than UCM. Body state variable $x$ is a minimal space of variables required to specify task execution. For each different subject, the GEM remain the same no matter the subject is a one-degree freedom actuator or a human with thousands of degrees of freedom.

For goal-level fluctuation, the error of performance is defined as

$$e = f(x; p)$$

(2-6)
Figure 2-1 **Goal Equivalent Manifold (GEM) for treadmill walking task.** Red dots are experimental data collected from treadmill walking task. The blue straight line represents the goal equivalent manifold defined by Eq. 2-4 and Eq. 2-5. Null subspace has the same direction with the GEM. The orange line represents goal relevant subspace which is perpendicular to the GEM, and its intersection with GEM is the operating point $\mathbf{x}^*$ of the task.

When $\mathbf{x}$ is not landed on the GEM, we have $\mathbf{e} \neq \mathbf{0}$. To relate goal-level error to body-level fluctuation, we rewrite $\mathbf{x}$ as,

$$
\mathbf{x} = \mathbf{x}^* + \mathbf{u}
$$

where $\mathbf{x}^* = [L^*, T^*]^T \in \mathcal{G}$. Usually we use $\bar{\mathbf{x}} = \mathbf{x}^*$ in experimental practice, since we assume that, for skilled performance, the average of body state over multiple trials is close to the GEM. Observed body-level error is defined as $\mathbf{u} = [p, q]^T$ in Eq. 2-7.

Substitute Eq. 2-7 into Eq. 2-6 and ignore $\mathbf{p}$ to simplify the notation, we have,

$$
\mathbf{e} = f(\mathbf{x}^* + \mathbf{u})
$$

For small fluctuation, we can mathematically do a first order Taylor Series at $\mathbf{x}^*$. Then we yield a linearized function to approximate goal-level error based on observed body-level deviation.
\[ e = f(x^* + u) \approx f(x^*) + Bu = Bu \] (2-9)

where matrix \( B \) is defined as body-goal variability matrix. It equals to the partial derivatives \( \frac{\partial f}{\partial x} \) at \( x^* \) (in our case, \( B = \left[ \frac{\partial f}{\partial l} \frac{\partial f}{\partial r} \right] \) at \( \bar{x} \)). By definition, matrix \( B \) has dimension of \( D_g \times D_b \). Here, we define two subspaces. The first one is goal relevant subspace, denoted by \( \mathcal{R} \) with dimension as same as \( D_g \). The fluctuation in this subspace will result in a goal-level error. The second one is null subspace, denoted by \( \mathcal{N} \) with dimension of \((D_b - D_g)\). In contrast, fluctuation in \( \mathcal{N} \) will not have any effects on performance. Note in Eq. 2-9, we have \( \mathcal{N} = \{ u|Bu = 0 \} \). These two subspaces can be visualized as two linear subspaces which are respectively tangent and perpendicular to the GEM.

2.2 Modeling Inter-Stride Fluctuations

In the last section, we reviewed the GEM of a dynamic system and related goal-level error to the observation, which represents body-level deviation. Since there are no assumptions of control mechanism underlay the development of goal equivalent manifold, GEM widely exists in dynamic systems, even for the one is uncontrolled. In this section, we will review how to abstract dynamical properties of a system from its experimental observations.

Observations obtained from experiments usually consist of a series of body state variables, \( \{ x_k \}_{k=1}^N \), while the task has been carried out through \( N \) strides. We consider the inter-stride behavior is governed by:

\[ x_{k+1} = x_k + G(I + N_k)c(x_k) + \nu_k \] (2-10)

where \( c(x_k) \) is the controller depending only on current body state; \( N_k \) is a matrix showing signal-dependent noise in the motor output; \( \nu_k \) represents all the other noise, like environmental source or perceptual fluctuations. \( G \) denotes the diagonal matrix of gains. In the absence of noise, the body
state variable will remain on the GEM with constant task performance. On the other hand, if the system is uncontrolled with Gaussian white noise, the fluctuations approach the behavior of a random walk.

For skilled performance, we can linearize Eq. 2-10 at a certain operation point $x^*$ based on assumptions that both deviations from GEM and noise inputs are small to get a first-order AR equation:

$$u_{k+1} = Au_k + v_k$$ (2-11)

in which $A = I + GJ$, where $J = \partial c / \partial x$ is the Jacobian of $c$ at $x^*$. Note that, the eigenvectors and eigenvalues of matrix $A$ indicate the local stability properties of the system. When the magnitudes of all the eigenvalues $\lambda_i$ are less than one at the operation point, it is defined as asymptotically stable. In this case, small perturbation will fade to zero with absence of noise over several strides. If there is one or even more eigenvalues $|\lambda_i| > 1$, fluctuations will be amplified over trials which makes the operation point unstable. When $|\lambda_i| = 1$, system is neutrally stable in the state space.

The previous study (Cusumano, Mahoney, & Dingwell, 2014) shows that in human movements, the unstable operation points usually will not be observed. And for a typical skilled task with $D_b = 2$ and $D_g = 1$, there will be an eigenvalue $\lambda_1$ which is slightly bigger than zero. Fluctuations along its corresponding eigenvector will be diminished quickly so that we recognize it as strongly stable. On contrast, the other eigenvalue is less than but close to one, which allows fluctuations carry over multiple trails. The corresponding eigenvector is then recognized as weakly stable. The above concepts are visualized in Fig. 2-2, notice that generally the weakly stable does not equal to the null subspace of GEM, but there is a small angle between them.
Figure 2-2 A typical MIP controller with weak POP term in body state space (Cusumano & Dingwell, 2013) with $D_b = 2$ and $D_g = 1$. We have $0 < |\lambda_1| < 1$ and $|\lambda_2|$ close to 1, which means it is strongly stable along the direction of $\lambda_1$, meanwhile, weakly stable along $\lambda_2$.

2.3 Building AR Model and Model Selection

The first-order AR model in the last section has been used to study gait of treadmill walking in previous studies. It shows for both young and old group, the dynamic properties share the same geometric structures and they both strictly regulate strongly stable direction. However, it is generally believed that the old has larger variability in movement, which is also tested to be true in our study.

Inspired by analyzing the variability from strongly and weakly stable directions, we made two experimental hypotheses regarding two subspaces trying to explain what happened to the controller of the dynamic system under the effect of aging. They are stated as follow:
**H1** Along goal relevant subspace $\mathcal{R}$, the controller does not significantly change while aging. The controller has to regulate goal relevant error closely so that the task can be guaranteed to execute well.

**H2** Along null subspace $\mathcal{N}$, the controller has higher dimensions of state space that allows information of fluctuations carry over more strides when age goes up. Or from another perspective, we can say the old uses more “history” in the movement to decide the next move.

Above hypotheses make the fluctuations in null subspace as the main contributor to the performance variability. Since the controllers of older subjects are more complex, it makes the movement of the old “noisier”.

To validate above hypothesis, we built a set of vector autoregressive (VAR) models with multiple orders by doing the linear regression. If the H1 is true, we are expecting models with lower model orders are preferred to fit experimental data, and preferred models are similar in two age groups. For validation of H2, higher order models are preferred and the preferred lag orders of the old group should be significantly higher.

Here we demonstrate the processes to develop models with the different order of lag, and how to use appropriate criteria to pick the fittest model. To apply Eq. 2-11, we convert the experimental data set to have form of $\{u_0, u_1, u_2 \cdots u_{k+1}\}$. Given the order of the AR model $m$, which indicates how much “history” is involved in the process of predicting “future”, the simplest way to predict $u_{k+1}$ would be linear:

$$u_{k+1} = \sum_{n=0}^{m-1} A_n u_{k-n} + \nu_{k-n} = A u_{k,m} + \nu$$

(2-12)
where
\[
\mathbf{u}_{k,m} = \begin{bmatrix} p_k & q_k \end{bmatrix} \begin{bmatrix} p_{k-1} & q_{k-1} & \cdots & p_{k-m+1} & q_{k-m+1} \end{bmatrix}^T
\]
(2-13)

\[
\begin{bmatrix} p_k \\ q_k \end{bmatrix} = \begin{bmatrix} L_k \\ T_k \end{bmatrix} - \begin{bmatrix} L^* \\ T^* \end{bmatrix}
\]
(2-14)

\[
\mathbf{A} = \begin{bmatrix} A_0 & A_1 & \cdots & A_{m-1} \end{bmatrix}
\]
(2-15)

As an intrinsic part of the system, noise is introduced as term \( \mathbf{v} \) into the equation: here we set \( \mathbf{v} \) to be an independent identically distributed Gaussian random noise. However, the experimental noise is not necessarily matching our preset of \( \mathbf{v} \), if \( \mathbf{v} \) has been tested to be non-Gaussian, which is not the case in our study, then a low pass filtering process shall be applied and an Auto Regression Moving Average (ARMA) model shall be used instead.

For a given lag \( m \), the only unknown parameter of the AR model is the coefficient matrix \( \mathbf{A} \). To determine each element in \( \mathbf{A} \), linear regression are used to minimize the residual, as shown in Eq. 2-16. \( \mathbf{v} \) is neglected when doing the regression.

\[
e^2 = \sum_{k=0}^{m-1} \left| \mathbf{u}_{k+1} - \mathbf{A} \mathbf{u}_{k,m} \right|^2
\]
(2-16)

In this thesis, we perform the calculation processes on MATLAB. To set up the linear regression, we first rewrite the data set to have the form of Eq. 2-13,

\[
\mathbf{u}_{k,m} = \begin{bmatrix} p_k \\ q_k \\ p_{k-1} \\ q_{k-1} \\ \vdots \\ p_{k-m+1} \\ q_{k-m+1} \end{bmatrix}
\]
(2-17)

and set

\[
\mathbf{K} = \begin{bmatrix} A_0 & A_1 & \cdots & A_{m-1} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix}
\]
(2-18)
Notice here the size of $\mathbf{u}_{k,m}$ is $2m \times 1$, and the size of $\mathbf{A}$ is $2m \times 2m$. In the auto regression model, the coefficient matrix $\mathbf{K}$ of certain $m$ remains constant and will not change when $k$ increases.

Then, we set

\[
\mathbf{u}_{r_{hs}} = \begin{bmatrix} u_{N-1,m} & u_{N-2,m} & \cdots & u_{m,m} \end{bmatrix}
\]

\[
= \begin{bmatrix} p_{N-1} & p_{N-2} & \cdots & p_m \\ q_{N-1} & q_{N-2} & \cdots & q_m \\ p_{N-2} & p_{N-3} & \cdots & p_{m-1} \\ q_{N-2} & q_{N-3} & \cdots & q_{m-1} \\ \vdots \\ p_{N-m} & p_{N-m-1} & \cdots & p_1 \\ q_{N-m} & q_{N-m-1} & \cdots & q_1 \end{bmatrix}
\]

where $N$ denotes the size of collected experimental data. $\mathbf{u}_{lhs}$ is $\mathbf{u}_{r_{hs}}$ shifted by one. Learnt from the setting of our AR model, we find

\[
\mathbf{u}_{lhs} = \begin{bmatrix} u_{N,m} & u_{N-1,m} & \cdots & u_{m+1,m} \end{bmatrix}
\]

\[
= \begin{bmatrix} p_N & p_{N-1} & \cdots & p_{m+1} \\ q_N & q_{N-1} & \cdots & q_{m+1} \\ p_{N-1} & p_{N-2} & \cdots & p_m \\ q_{N-1} & q_{N-2} & \cdots & q_m \\ \vdots \\ p_{N-m+1} & p_{N-m} & \cdots & p_2 \\ q_{N-m+1} & q_{N-m} & \cdots & q_2 \end{bmatrix}
\]

Now we can simply find the expression of $\mathbf{K}$ to get $\mathbf{A}$ by taking the dot product between $\mathbf{u}_{r_{hs}}$ and the generalized inverse (known as well as pseudoinverse or Moore-Penrose inverse (Moore, 1920)) of $\mathbf{u}_{lhs}$ on MATLAB using command `pinv`.

After calculation of $\mathbf{A}$ of each data set for given model order $m$, we have a set of AR models. Be able to find the model that fits experimental data the best is the key point of this study.
We picked two well-known information criteria, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), to determine the quality of each model. The better the model fits, the lower the value of its AIC or BIC will be. Thus, the model with the lowest value of AIC or BIC is selected and its model order is recorded.

In our study, for each trial, we tested 10 models with model orders ranged from 1 to 10. For the model with order higher than 10, its value of AIC or BIC is significantly larger than models with the lower order because of the penalty of overfitting in both criteria is getting stronger.

### 2.4 Significance Test

In our study, the results usually follow discrete non-normal distribution so that we are not able to use some widely used parametric test, like t-test. Permutation test method, a non-parametric model, has been applied to evaluate significance in this thesis.

Permutation test is a non-parametric statistical significance test which rearranges obtained data points to get all possible combinations, then calculate distribution under the null hypothesis (Collingridge, 2013). In our study, we first find and record the difference of interested statistic properties, such as mean, median, etc., between two interested groups. Then, we mix all the data together regardless their original group, and randomly reassign them into two new groups and finds the new difference. By repeating above reassignment processes for finite times R, we get a distribution of the difference between randomized data. The p-value can be evaluated by finding the ratio of cases that a larger or smaller difference is observed in the new group to the randomization times R.

Usually, the interested statistic properties are picked by following rules: if the data are approximately Normal, then use the difference of two means; if the data are symmetric but outliers are present, then use the difference of two trimmed means; if the data are not symmetric,
then use the difference of two medians. We picked median as interested statics to test the significance of variability, and mean as interested statics to test model order.

### 2.5 Data Processing in Software Package

The time series model set up and model selection are performed using MATLAB (MathWorks, Natick, MA). In MATLAB, since our data type of time series model is vector, we choose to use VARMAX model related functions in Econometrics Toolbox for testing.

Firstly, we pretreat the raw data collected from the experiment by transforming it to variability $\mathbf{u}$, and constructing six data groups: Overall_Young, Overall_Old, Normal_Young, Normal_Old, Tangential_Young and Tangential_Old. Here, “Overall” in above groups means we did not distinguish data collected under different walking speed. “Normal” and “Tangential” denotes projections in goal relevant subspace and null subspace.

Next, for each data group, we get $\mathbf{u}_{ths}$ and $\mathbf{u}_{rhs}$ and use command `pinv` to calculate matrix $\mathbf{A}$ for given lag $m$. Then function `vgxset` is used to set VARMAX model parameters. To meet the setup of our model, we let order of moving-average (MA) to be zero, and order of AR equals to $m$. Two input arguments are required for command `aicbic` to calculate the value of AIC and BIC. The likelihood value $\hat{L}$ is calculated by `vgxvarx`, and the free parameter number $k$ in AR model is counted by `vgxcount`. See appendix for related code.

As mentioned in last section, we set $m$ from one to ten, and repeat above processes till all the AIC and BIC values for each $m$ have been recorded. After selecting preferred model, statistical significance tests are performed using Minitab (Minitab Inc. State College, PA) with permutation method. For selected model orders, the distribution is asymmetrical so that we record the difference of means of paired groups, like Overall_Young and Overall_Old. We set $R = 10000$ and run
randomization R times. The differences of new groups are recorded and plotted in the histogram together with a red reference line showing the difference between the original paired groups. The larger the area of bars on the right of the reference line, the bigger the p-value is. See Chapter three for results of permutation tests.

2.6 Subjects and Experiment Setup

Each of seventeen young and old healthy adults was chosen as subjects. The young group is consist of 12 males and 5 females, age ranged from 18 to 28, height 1.73±0.09 m, and weight 71.11±9.86 kg. The old group is consist of 11 males and 6 females, age ranged from 66 to 78. All the participants do not have any history of the orthopedic problem, lower extremity injury, visible gait anomalies, or under treatment which may affect their performance.

The treadmill used in the experiment is the Desmo S model made by Woodway USA, Waukesha WI. For the walking task, subjects were asked to walk on the even level with comfortable shoes and no restriction for arm swing. Subjects reported their preferred walking speed (PWS) based on the average of their highest limit speed and lowest speed. Subjects completed two five-minute walking trials at each of five speed (80, 90, 100, 110, 120% of PWS). At least two minutes rest was given between trials. Due to the technical reason, there are four subjects missing result for one trial.
Chapter 3

Results

In this chapter, we show the results from four aspects by comparing interested characteristics of two age groups. We plot the histogram to show the distribution of preferred model order, followed by permutation distribution showing significance between paired groups. The confidence level of the significance test is set to be 95% (p-value ≤ 0.05), that is saying we claim difference exists when the area of bars on the right of the red reference line is less than 5% to the area of all bars.

3.1 Variability of Young and Old Groups

We first validate the general knowledge by testing if the old participates have larger variability than the young. Fig. 3-1 shows the boxplots of inter-stride variability of two age groups and the corresponding projections in two subspaces. We see that in all cases the old group does have a larger variability. In aggregate, across all participants in young group, we find $\text{median}_{\text{young}} = 0.01935$, and a significantly higher $\text{median}_{\text{old}} = 0.02445$ is found in the old group. Similar results are found in the two subspaces, the old group has a roughly 45% higher level of variability in both subspaces. We note that the variability in goal relevant subspace is smaller than it in null space, it is because that the orientation of weakly stable is strictly regulated to be near the null subspace, whereas the orientation of strongly stable is close to but not equal to goal relevant subspace. Above findings supports the general knowledge while pointing out variability in the null subspace may be a main contributor to the overall variability.
Figure 3-1 Variability of Young Participants and Old Participants. (A) Boxplot of variability medians of both young (median = 0.01935) and old group (median = 0.02445). Old group has a significantly (p-value < 0.05) higher variability than the young group. (B) Boxplot of two group showing variability respected to normal and tangential direction. The variability of normal direction (median = 0.00817) is 43% less than it of tangential direction (median = 0.01441) for the young group, and the variability of normal direction (median = 0.01004) is 45% less than it of tangential direction (median = 0.01833) for the old group.

3.2 Results regarding overall group

In Fig. 3-2, we find the most probable model orders for both groups are no larger than three. Moreover, AIC shows the third order model best describes overall treadmill walking task no matter of age, while BIC shows young subjects prefer the first order model and old subjects prefer the second order model.

When we go further to test the statistical significance of the difference of means between two groups, two criteria do not show consistency. Models picked by AIC do not show significant difference between young group (mean = 3.5181) and old group (mean = 3.6071) with a p-value equals to 0.3033± 0.0015, while models selected by BIC strongly support that old group (mean = 1.7619) prefers higher order model than young group (mean = 1.5663) by showing the significance with a p-value equals to 0.0044± 0.00002. Moreover, we find BIC generally prefers lower model order compared to AIC, since penalty of overfitting in BIC is larger than in AIC.
Figure 3-2 **Treadmill walking model order selection and significance test.** (A) Distributions of preferred model orders picked by AIC. Young (Red) and old groups (Blue) have slightly different distribution of its preferred model. The third order model is the most probable model for both young and old groups (B) Distributions of preferred model orders picked by BIC. The highest order of preferred models is four compared to nine for AIC picked models since BIC has more penalty for overfitting. We can see that young subjects prefer the first order model and old subjects prefer the second order model (C) Permutation distribution of the difference of means. The randomization sampling size $R=10000$, the difference of means of observations is 0.008907, shown as the red reference line. In the permutation distribution, there are 3033 samples, which marked as mesh red bar, have larger difference of means than observations. Thus, in AIC picked models, difference of model orders between young group and old group is not statistically significant ($p = 0.3033 \pm 0.0015$). (D) Permutation test for BIC shows different result compared to AIC. Under the same randomization conditions, there are only 44 samples have larger difference of means, which shows a great statistical significance ($p = 0.0044 \pm 0.00002$) of difference between two groups.

### 3.3 Results regarding Normal and Tangential Directions

To validate the two hypothesis in chapter three, we expect an insignificant difference exists in goal relevant subspace between young and old to support H1, while a significant difference exists in null subspace to support H2.
Figure 3-3  **Treadmill walking model order selection and significance test in goal relevant subspace.** (A) Distributions of preferred model orders of projection on goal relevant subspace, picked by AIC. Young (Red) and old groups (Blue) have similar distribution of the best fit model orders. (B) Distributions picked by BIC shows that for both young and old group, in goal relevant subspace, they all tend to prefer a first order model. (C) In the permutation distribution of AIC selected model orders, there are 1256 samples have larger difference of means while randomization sampling size R=10000. Difference of model orders between young group and old group is not statistically significant (p = 0.1256 ± 0.0006). This finding supports the Hypothesis **H1**. (D) Permutation test for BIC agrees with the finding of AIC. It is statistically insignificant (p = 0.0551 ± 0.0003) to claim difference exists in two age groups.

For model orders of projections of variability on goal relevant subspace, permutation test gives significant level larger than 0.05 for models orders selected by both AIC and BIC (p-value = 0.1256 ± 0.0006 for AIC and p-value = 0.0551 ± 0.0003 for BIC). Thus, we cannot claim there is difference of means between young and old groups (mean_young_AIC = 1.9684, mean_old_AIC = 2.1634, mean_young_BIC = 1.1506, mean_old_BIC = 1.2679). This finding supports **H1** by proving that subjects in the old group regulate goal relevant subspace just as well as subjects in the young group do.
For model orders of projections of variability on null subspace, permutation tests for AIC and BIC selected model orders give different results. Test for AIC supports H2 by finding p-value = 0.0396 ± 0.0002, and claiming subjects in old group allow variability carry over more trails than the young does in null subspace(mean_young_AIC = 4.8433, mean_old_AIC = 5.2857). On the contrary, test for BIC rejects H2 with a p-value = 0.1122 ± 0.0006 claiming difference between two groups is not significant (mean_young_BIC = 2.7636, mean_old_BIC = 2.9345).
Figure 3-5 **Dynamical geometry structure with respect to GEM of young group.** (A) Distributions of preferred model orders for young group, picked by AIC, in goal relevant subspace and null subspace are placed side by side. Model orders in goal relevant subspace (orange) have lower orders than those in null subspace (plum). (B) Distributions of preferred model orders picked by BIC show similar structure with lower preferred orders compared to AIC picked models. (C) Permutation distribution of difference of means. The randomization sampling size R=10000, and zero sample has larger difference of means. Difference of model orders between young group and old group, in AIC picked models, is statistically significant (p=0.0001 ± 5E-7). There is no red reference line in the chart since it is beyond the scale. (D) Permutation test for BIC shows exact same result with AIC. Zero sample has larger difference of means, which shows a great statistical significance (p=0.0001 ± 5E-7) of difference between two groups.

### 3.4 Results regarding Dynamical Geometry respected to GEM

In previous work, the study of GEM shows the eigenvectors and eigenvalues of matrix A in Eq. 2-11 are indicators of weakly stable and strongly stable directions. The direction of goal relevant subspace is closely near the direction of strongly stable, the magnitude of fluctuation along normal direction are corrected fiercely. For null subspace, usually there is a small angle between
Figure 3-6 **Dynamical geometry structure with respect to GEM of old group.** (A) Distributions of preferred model orders for old group, picked by AIC, in goal relevant subspace and null subspace are placed side by side. Model orders in goal relevant subspace (orange) have lower orders than those in null subspace (plum). The structure of old group is similar to young group in Fig. 3-5. (B) Distributions of preferred model orders picked by BIC show similar structure with lower preferred orders compared to AIC picked models. (C) Difference of model orders between young group and old group is statistically significant (p=0.0001 ± 5E-7). No difference larger than 1 is got during randomization while difference of 3.12232 is observed in experiment. (D) Permutation test for BIC shows a great statistical significance (p=0.0001 ± 5E-7) of difference between two groups.

its direction and weakly stable direction, as shown in Fig. 2-2. The magnitude of fluctuation along tangential direction decreases much slower compared to the normal direction.

In our study, distributions of the best fit orders along two directions are significantly different (Fig. 3-5 and 3-6) no matter of age. For orders selected by AIC and BIC, there is no random combination has larger difference of means than observations, which leads to a p-value of 0.0001 ± 5E-7. This finding indicates a higher dimension of state space is required to accurately model fluctuation in null subspace.
Chapter 4

Conclusion and Discussion

In this thesis, we first reviewed the concept of the Goal Equivalent Manifold (GEM) data analysis framework, which is a previously-developed approach to study inter-stride error correction and to predict both the temporal and geometric structure of variability near a goal equivalent manifold for skilled performance. Then, we introduced the idea of time series analysis into the framework to model experimental data with lags, so that we might be able to locate the origin of extra variability in old people, especially given the finding of previous studies of GEM that both young and old people share a similar geometrical structure to their variability. Model order selection within the GEM framework allows us to evaluate the quality of model fits for different lags. We applied the time series analysis to treadmill walking data to explore best fit model order which can describe the inter-stride error correction processes, and to answer how preferred model order changes with age. We paired up the data for testing under several purposes, such as comparing differences exist between two age groups or two subspaces. Finally, in this chapter, we will conclude our findings and discuss them based on the understanding gained from our analysis.

Old people are generally believed to have “noisier” movement than the young. In treadmill walking task, we have justified the general knowledge by showing the variability of the old group is significantly higher than young subjects. When we try to address the cause of above observations using time series analysis, two information criteria used in our study showed inconsistent results regarding the difference of preferred model order in the two groups. Model orders selected by AIC indicated no significant difference exists between the two groups, whereas model orders selected by BIC supported that young participants have a lower order model and use less “history” from previous strides for executing the next move. Accordingly, we cannot claim the larger variability
of old people may be caused by a more complex control model with higher dimension of state space.

In the previous study, a “GEM aware” controller, which exhibits weak stability along the GEM and strong stability transverse to it, have been used to describe how humans accurately carry out repeated strides when walking on a treadmill. The above idea inspired and led us to two hypothesis that aimed to explain the cause of extra variability in the old group to the geometric structure respecting to the GEM. Note that instead of analysis weakly stable and strongly stable direction, we projected the variability onto goal relevant subspace and null subspace.

In goal relevant subspace \( \mathcal{R} \), both AIC and BIC showed that a low order model (typically a first or second order model) best fit the observation in goal relevant subspace. This supported that controller does not significantly change in goal relevant subspace while aging so that it can regulate goal relevant error closely to guarantee the task execution, that is saying \( \textbf{H1} \) is accepted.

However, in null subspace \( \mathcal{N} \), another inconsistent finding between methods of AIC and BIC showed up when validating hypothesis \( \textbf{H2} \): old people use more information from previous strides to execute the next one. We expected to ascribe the extra variability in the old group to more “complex” regulation in the null subspace since both young and old groups regulate fluctuations well in the perpendicular goal relevant subspace. The AIC method supported \( \textbf{H2} \) by showing the old group has a significant higher model order in the null subspace, as we expected, whereas BIC method rejected \( \textbf{H2} \) by showing the difference is not significant.

We also found a significant tendency that a lower order model better describes the underlying processes in goal relevant subspace, while a higher order model is preferred in null subspace. This finding indirectly supported the investigations of the geometric structure in previous studies regarding weakly stable and strongly stable. In our study, a preferred low order model in goal relevant subspace indicates less previous information has been taken into account. Meanwhile, previous works show that fluctuations along strongly stable are eliminated more rapidly. Similarly,
high order model are preferred in null subspace and slow correction processes are found in weakly stable direction.

We have seen that to describe the motor control system in human movement, that is, the system that regulates motor variable, model with the lag order under ten is enough to model underlying fluctuations. In particular, both AIC and BIC method demonstrated that the first order model is the most common model for both young and old group in goal relevant subspace. Considering the complexity of higher order model adds marginal additional insight to the dynamic system, and the first order autoregressive model used in previous GEM studies have already successfully generalized multiple tasks, such as reaching and walking. When modeling future problems of movement dynamics, one has to balance between the accuracy contributed by higher model order with the added complexity.

These are still two inconsistent findings left in this thesis. Two major improvements can be done in the future studies. In our thesis, we project the variability onto two subspace instead of weakly and strongly stable directions as suggested by corresponding eigenvectors. Since directions between these two systems do not perfectly equal to each other, studies of projections onto weakly and strongly stable directions could potentially be the key to understand the difference between young and old groups. The second improvement could be done by introducing more information criteria, like Deviance information criterion and Hannan-Quinn information criterion, or improve and inspect the regular AIC and BIC more carefully, such as adopting relative likelihood method when deciding the best fit model order.
References


Appendix

MATLAB Software for Data Analysis

Module for calculating Body-Goal variability matrix

```matlab
%% Input: TrD (True Delay), E (Original data sheet)
%% Output: ACo (Body-Goal variability matrix)
%% Comments: Null

for TrD = 1:10

E1=zeros(TrD*2,(length(E)-TrD)); % Initializing Matrix E1 to restore
data of prior strides
corresponding to True Delay TrD
E2=E1; % Initializing Matrix E2 to restore
data of successive strides
corresponding to TrD

for k = 1:TrD % Loop to get Matrix A for each TrD
    E1(2*TrD-2*k+1:2*TrD-2*k+2,:)=E(k:(length(E)-TrD-1)+k,:); 
    E2(2*TrD-2*k+1:2*TrD-2*k+2,:)=E(k+1:(length(E)-TrD)+k,:); 
end

A=E2*pinv(E1); % pseudoinverse

if TrD>1
    A(3:TrD*2,1:TrD*2-2)=eye(TrD*2-2); 
    A(3:end,TrD*2-1:end)=zeros(2*TrD-2,2); 
end

ACo=A(1:2,:); % ACo here is Matrix A

end
```
Module for finding AIC and BIC for each Model Order

```matlab
%% Input: TrD (Model Order), ACo (Body-Goal variability matrix)
%% Output: aic (Matrix restore AIC), bic (Matrix restore BIC)
%% Comments: Null

for i=1:TrD
    ACoM(i)=ACo(1:2,TrD*2-1:TrD*2); % Break down Matrix A into several 2x2 matrix for each order of lag
end

Spec=vgxset('n',2,'nAR',TrD,'AR',ACoM); % Set AR(TrD) model

[SpecE,ESE,LogL,W]=vgxvarx(Spec,E); % Estimating Likelihood
NumP=vgxcount(SpecE); % Counting # of parameters
[a,b]=aicbic(LogL,NumP,length(E)); % Estimating AIC and BIC

aic(TrD)=a; % Restoring AIC
bic(TrD)=b; % Restoring BIC
end```