Using Switching State-Space Models to Interpret Musical Dynamics

Robert Granger *
Department of Statistics, Indiana University

March 10, 2020

Abstract

The text of your abstract. 200 or fewer words.

Keywords: 3 to 6 keywords, that do not appear in the title

*The authors gratefully acknowledge . . .

1 Introduction

Things to write about in this section

1) Classifying music performances to sort on personal preference

2) Smoothing performances to sound better from Midi data

3) Using State-Switching Model

With a piece of classical music, directions such as

2 The Switching State-Space Model

State-Space models are commonly used to model time series observations in the presence of perceived hidden, continuous states. As a result of the state-space framework, the observations are viewed as independent conditional on the hidden states whereas these hidden states will follow a vector autoregressive process. Adding assumptions of linearity and Gaussian error produces the following model commonly referred to as the general linear Gaussian state-space model Durbin & Koopman (2012) takes the form

$$\begin{aligned} y_t &= C_t + D_t x_t + \epsilon_t, \quad \epsilon_t \sim N(0, G_t) \\ x_{t+1} &= A_t + B_t x_t + \eta_t, \quad \eta_t \sim N(0, H_t), \quad x_1 \sim N(x_0, P_0) \end{aligned} \tag{1}$$

where the first part of Equation 1 is known as the observation equation and the second part is known as the state equation. y_t is a vector of known observations, and x_t is a vector of the unobserved, continuous states at each time period, t. The observation error, ϵ_t, and the state equation error, η_t, are both assumed to be independent and identically distributed.

In the typical state-space framework, the matrices A_t, B_t, C_t, D_t, G_t, and H_t are allowed to vary across time but are known. If there are a finite number of perceived structures for these matrices, and the given structure is unknown at time, t, a switching state-space model can be used. This model assumes there are some underlying discrete states, s_i, that transition over time through a Markov process. Making this slight adjustment to Equation 1 yields the following model which will be used as a basic framework for modeling music dynamics.

$$
\begin{aligned}
y_t &= C_t(\Theta_t) + D_t(\Theta_t)x_t + \epsilon_t, \quad \epsilon_t \sim N(0, G_t) \\
x_{t+1} &= A_t(\Theta_t) + B_t(\Theta)x_t + \eta_t, \quad \eta_t \sim N(0, H_t), \quad x_1 \sim N(x_0, P_0)
\end{aligned} \tag{2}
$$

When it comes to musical dynamics, rarely do musicians attempt to play with the same dynamics or loudness throughout the entirety of a piece. While there may be dramatic changes occasionally throughout the performance, most of the time we expect the musician to steadily change the loudness from note to note. Furthermore, the musician may quietly or loudly play individual notes for emphasis. In order to model this behavior, we propose the following four discrete states:

s_1: The musician selects a new value for loudness.

s_2: The musician continues the dynamics in a steady way.

s_3: The musician plays a single note more loudly.

s_4: The musician plays a single note more softly.

The observation, y_t is the univariate loudness of the note at each time period, t. In order to allow the dynamics to progress steadily, we implement the framework of Gu & Raphael (2012) where the continuous hidden states follow a process that allows for piece-wise quadratic displays with

$$x_t = (x_t^0, x_t^1, x_t^2),$$

where x_t^0, x_t^1, and x_t^2 are the loudness, the first order difference, and the second order difference, respectively. Since we want to maintain this smooth progression even when the musician plays a single note more loudly or softly, states s_3 and s_4 are implemented by adding a constant in the observation equation as opposed to changing the state equation. Table 1 shows the parameter matrices for the four states.

The final part of the switching state-space model is designing the Markov process for transitioning between the discrete states. Figure 1 displays the structure of the transition between states. The first state, s_1, allows for the selection of a new loudness which then transitions into the smooth progresson state, s_2, with probability 1. When arriving in s_2, the next time period's discrete state can be any of the possible states including itself. If in

States	Parameter Matrices					
s_i	A	B	C	D	G	H
s_1	$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	0	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	σ_ϵ^2	$\begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$
s_2	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$	0	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	σ_ϵ^2	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
s_3	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$	μ_c	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	σ_ϵ^2	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
s_4	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$	$-\mu_c$	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	σ_ϵ^2	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Table 1: Parameter matrices for the switching state space model.

this progression, we move to a sudden loud note, s_3, or a sudden soft note, s_4, then we have the opportunity to continue the smooth progression, s_2, or start a new smooth progression, s_1. It is not permissible though to immediately then play another sudden loud or soft note. With four states, the transition matrix could potentially have 12 probabilities to estimate; however, because of restrictions placed on the possible transitions, only 5 probabilities need to be estimated.

3 Evaluating the Model

Things to include:

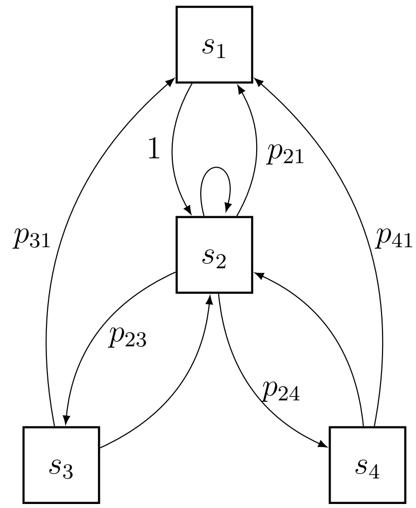

Figure 1: Transition diagram.

1) Discussion of state-space model algorithm with discrete particle filter

4 Conclusion

References

CHARM (2009), 'Centre for the History and Analysis of Recorded Music'.

Durbin, J. & Koopman, S. J. (2012), *Time Series Analysis by State Space Methods: Second Edition.*, Oxford University Press, Oxford.
URL: *http://grail.eblib.com.au/patron/FullRecord.aspx?p=991963*

Earis, A. (2007), 'An Algorithm to Extract Expressive Timing and Dynamics from Piano Recordings', *Musicae Scientiae* **11**(2), 155–182.
URL: *http://journals.sagepub.com/doi/10.1177/102986490701100202*

Fearnhead, P. & Clifford, P. (2003), 'On-line inference for hidden Markov models via particle filters', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**(4), 887–899.
URL: *http://doi.wiley.com/10.1111/1467-9868.00421*

Gu, Y. & Raphael, C. (2012), 'MODELING PIANO INTERPRETATION USING SWITCHING KALMAN FILTER', *13th International Society for Music Information Retrieval Conference* .

McDonald, D. J., McBride, M., Gu, Y. & Raphael, C. (2019), 'Markov-switching State Space Models for Uncovering Musical Interpretation', *arXiv:1907.06244 [stat]* .
URL: *http://arxiv.org/abs/1907.06244*

Wickham, H. (2016), *ggplot 2: Elegant graphics for data analysis*, 2nd edn, Springer.