HMM Extensions

RIIAA

MEETING ON ARTIFICIAL INTELLIGENCE AND ITS APPLICATIONS

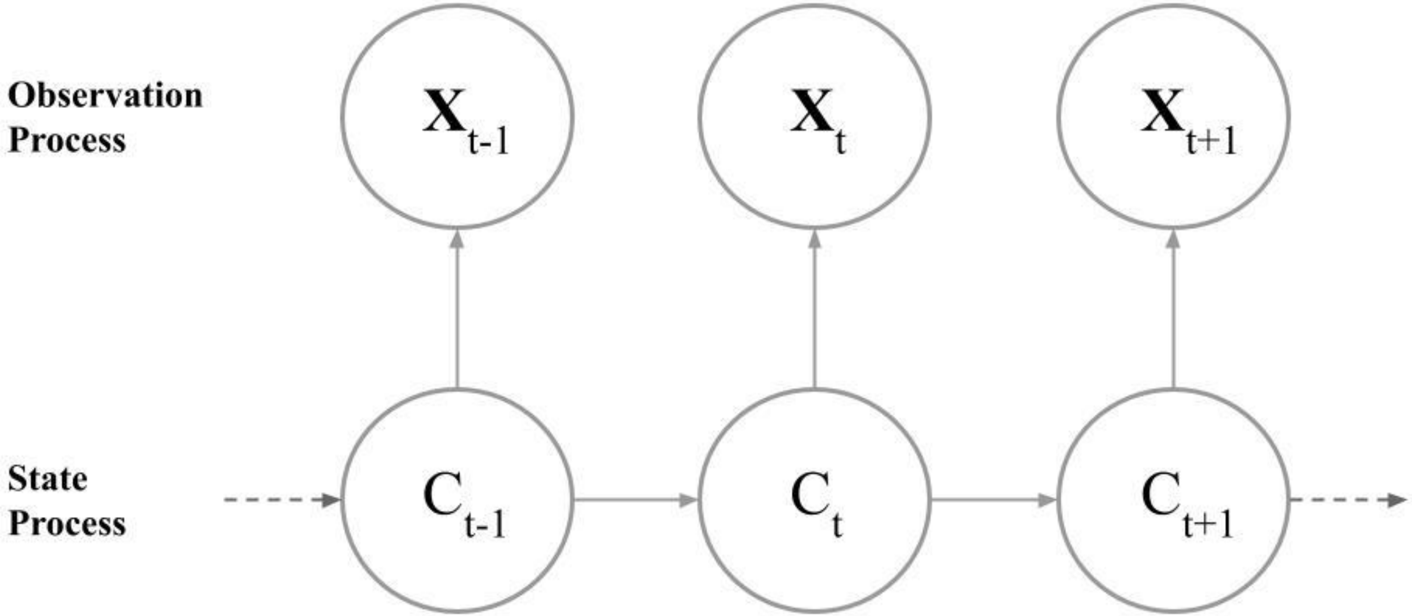
25-27 August 2021

HMMs Extensions

So far, we’ve covered the details of a basic HMM. But, much of the popularity in HMMs is due to the many ways it can be extended.

- Multivariate HMMs
- Covariates in the state or observation process
- Continuous-time HMMs
- Hidden semi-Markov models

Multivariate HMMs



Multivariate HMMs (cont.)

Our observation process is now **multivariate**: $\boldsymbol{X}_t = (X_{t,1}, \ldots, X_{t,K})$

The most common approach to assigning multivariate state-dependent distributions is to have $\boldsymbol{X}_t \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. But, we can also assign univariate distributions to $X_{t,k} \sim f(\cdot)$

Longitudinal independence: $\boldsymbol{X}_t \sim f(\cdot)$

Contemporaneous conditional independence: $\boldsymbol{X}_t \sim f(X_{t,1}) f(X_{t,2}) \cdots f(X_{t,K})$

Covariates

Introducing covariates into the model structure can be done in either the observation and/or state process. We can then estimate the effect that various factors had on the observations or the state-switching dynamics.

For example:

If data was collected across multiple animals, we might want to account for differences in the means of the state-dependent distributions according to different sizes. This can be done by $\mu_n(z_t) = \omega_0 + \omega_1 z_t$

Very commonly, the assumption of a single t.p.m. (a homogeneous Markov chain) over time is unrealistic. We can then incorporate covariates (including e.g. time of day) to construct an inhomogeneous Markov chain, $\mathbf{\Gamma}(z_t)$

Continuous-time HMMs

An HMM can also be formulated in continuous-time so that $\boldsymbol{X}_t$ does not have to occur ‘regularly’ over time. Two changes:

1. We are able to observe a ‘snapshot’ of the observation process
2. We estimate a *transition rate matrix* $\boldsymbol{Q}$, in order to construct the transition probability matrix. $\boldsymbol{\Gamma}_{\Delta_t} = e^{\boldsymbol{\Delta}_t \boldsymbol{Q}}$

Hidden semi-Markov models

Going back to the basic HMM structure, note that the implied amount of time that a state will be exhibited before the process switches is assumed to follow a **geometric distribution.**

This assumption can be quite unrealistic for some problems, but we can change the distribution so that $d_n \sim Poisson(\lambda_n)$ or any other valid distribution defined on the integers.

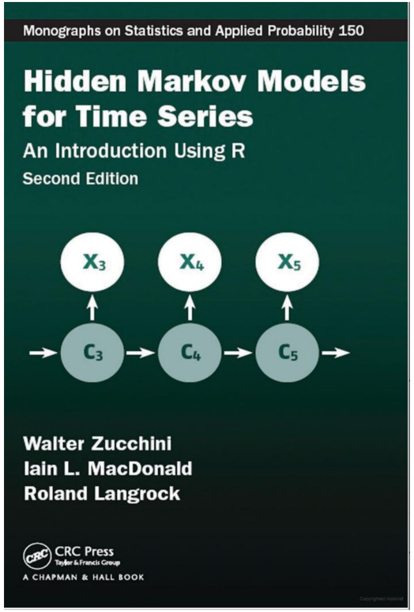
In summary....

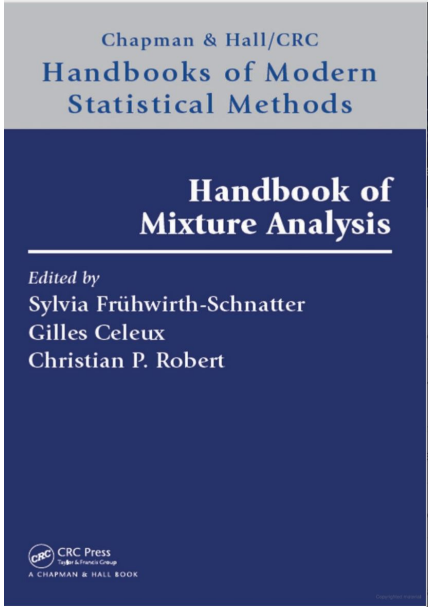
HMMs are extremely useful and highly adaptable to various problems across multiple domains.

They provide an **interpretable** framework for modeling, classification and pattern recognition.

Resources

[Pattern recognition and machine learning]





Contacto:

vianey.leosbarajas@utoronto.ca

sofi.ruizsuarez@gmail.com