Guided Tour of Machine Learning in Finance

Week 4: Reinforcement Learning

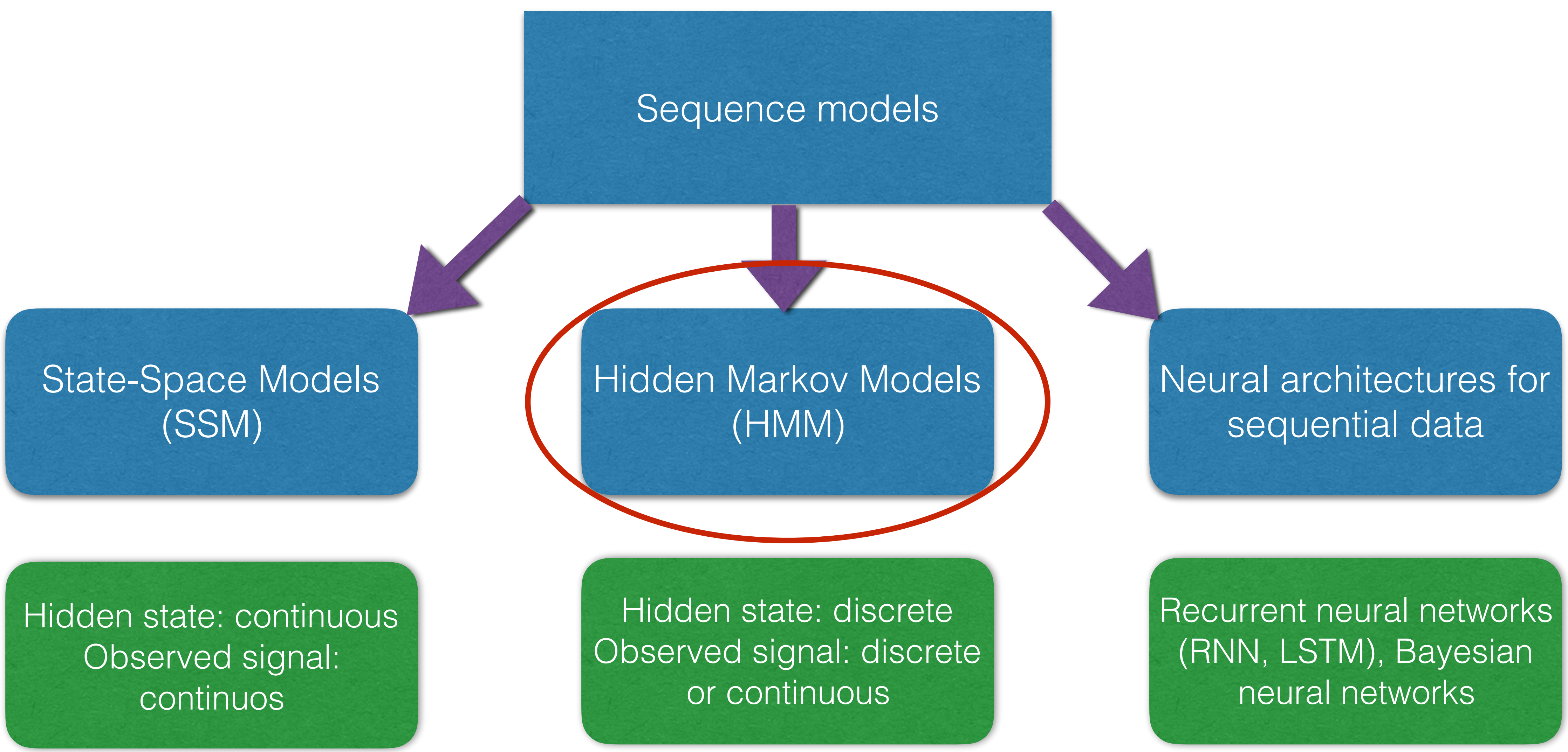
4-1-4-Hidden-Markov-Models

Igor Halperin

NYU Tandon School of Engineering, 2017

Sequence models

Parametric (SSM, HMM) vs non-parametric (neural) of sequence modeling with a hidden state $p(y \mid \mathbf{x})$



Hidden Markov Models (HMM)

The last: build a probabilistic model for the observable signal $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$

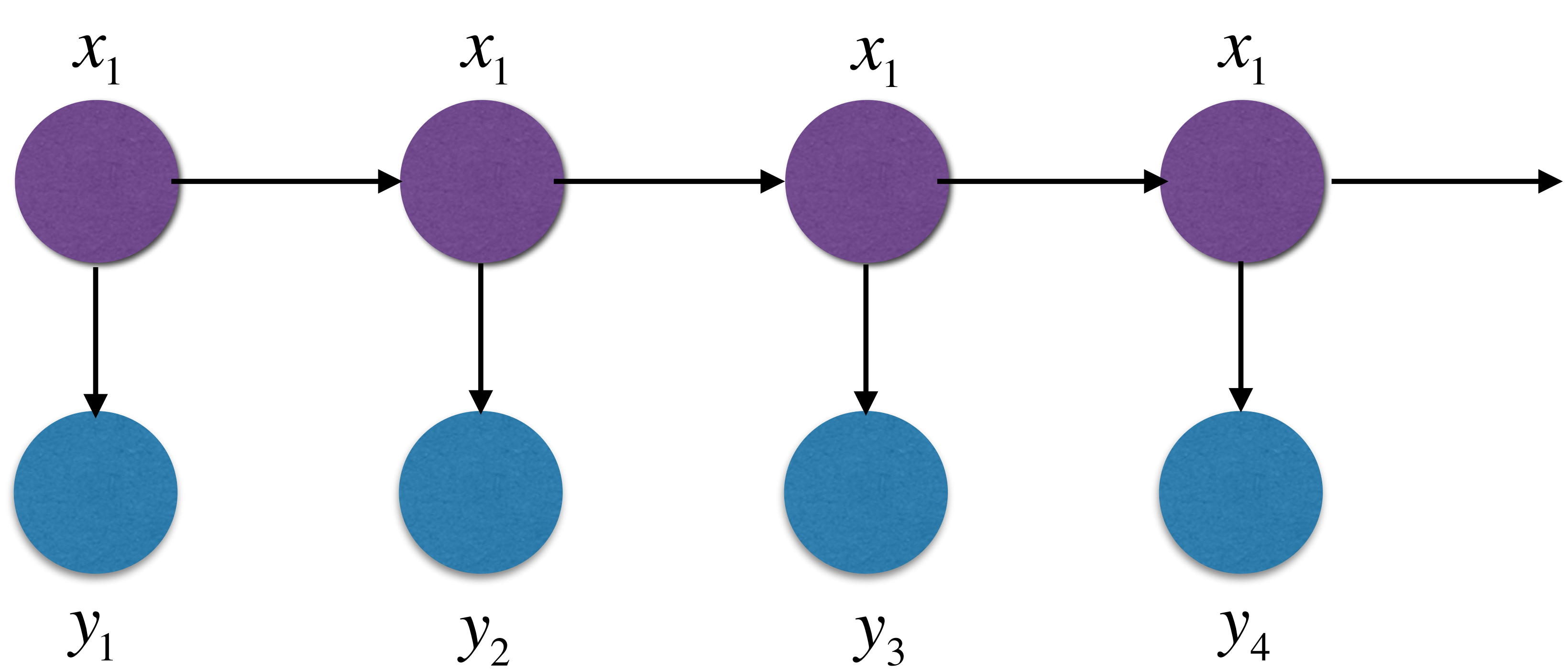
HMM = dynamic latent variables models with a **discrete** hidden state $x^{(t)}$

- The hidden state $x^{(t)}$ captures the dynamics of the system, filters noise out
- Used as a conditioning variable for predictions $p\left(y^{(t)} \mid y^{(t-1)}\right) \rightarrow p\left(y^{(t)} \mid x^{(t)}\right)$

The dynamics is first-order Markov in the hidden state:

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

- Hidden states x have first-order Markov dynamics encoded in $p(x_t \mid x_{t-1}, \theta)$
- Observations are generated from hidden states according to $p(y_t \mid x_t, \theta)$



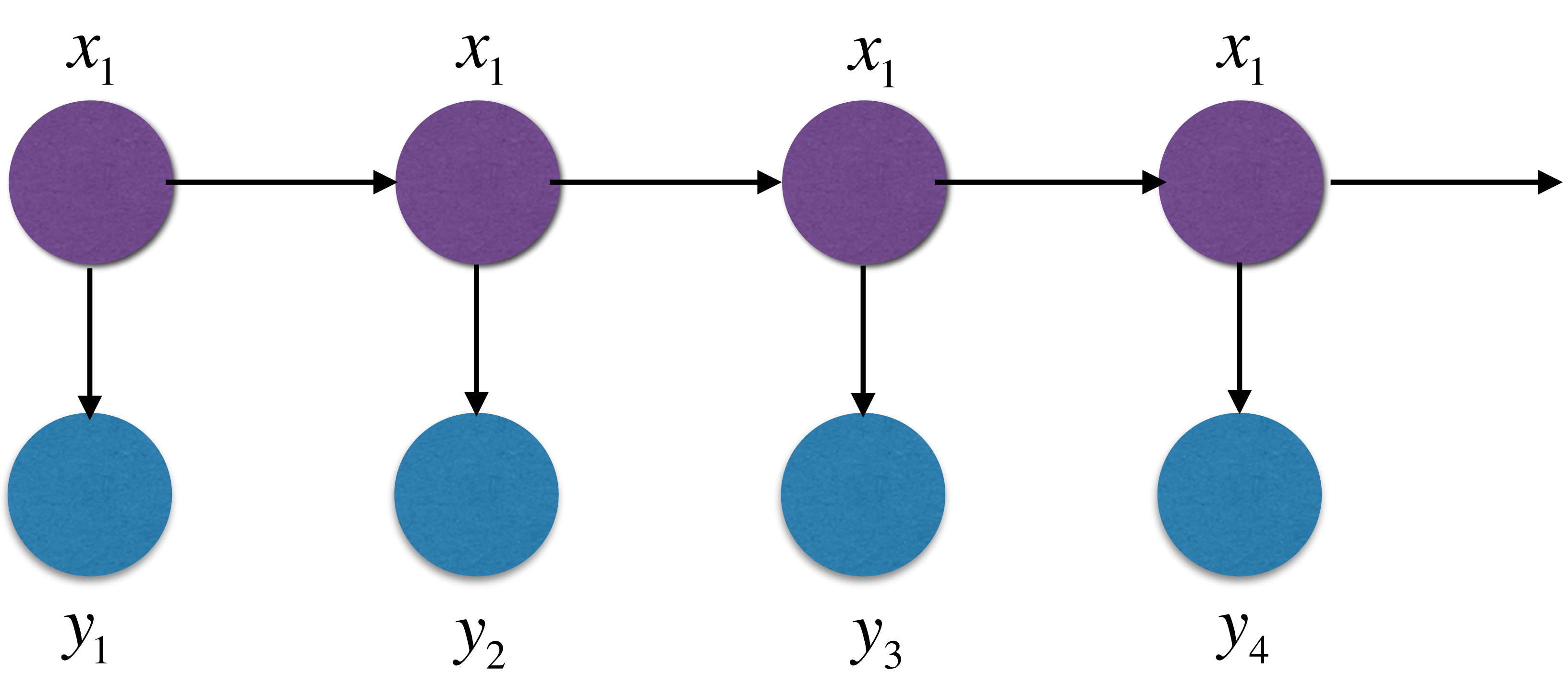
Observations in HMM

The dynamics is first-order Markov in the hidden state:

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

The observed signal can be:

- Discrete: $p(y_t \mid x_t, \theta)$ is a discrete distribution, e.g. a multinomial distribution
- Continuous: $p(y_t \mid x_t, \theta)$ is a continuous distribution, e.g. a Gaussian



Hidden Markov Models (HMM)

Observable N-dimensional data: $y_{1:T} = y^{(1)}, y^{(2)}, \ldots, y^{(T)}$
Hidden state sequence: $x_{1:T} = x^{(1)}, x^{(2)}, \ldots, x^{(T)}$

The dynamics is first-order Markov in the **discrete** hidden state that takes values in [1:K]:

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

- Hidden states x have first-order Markov dynamics encoded in $p(x_t \mid x_{t-1}, \theta)$
- Observations are generated from hidden states according to $p(y_t \mid x_t, \theta)$

Two computational problems:

Forecasting future value $p(y_{t+1} \mid x_{1:t}, y_t, \theta) = Inference + Learning$

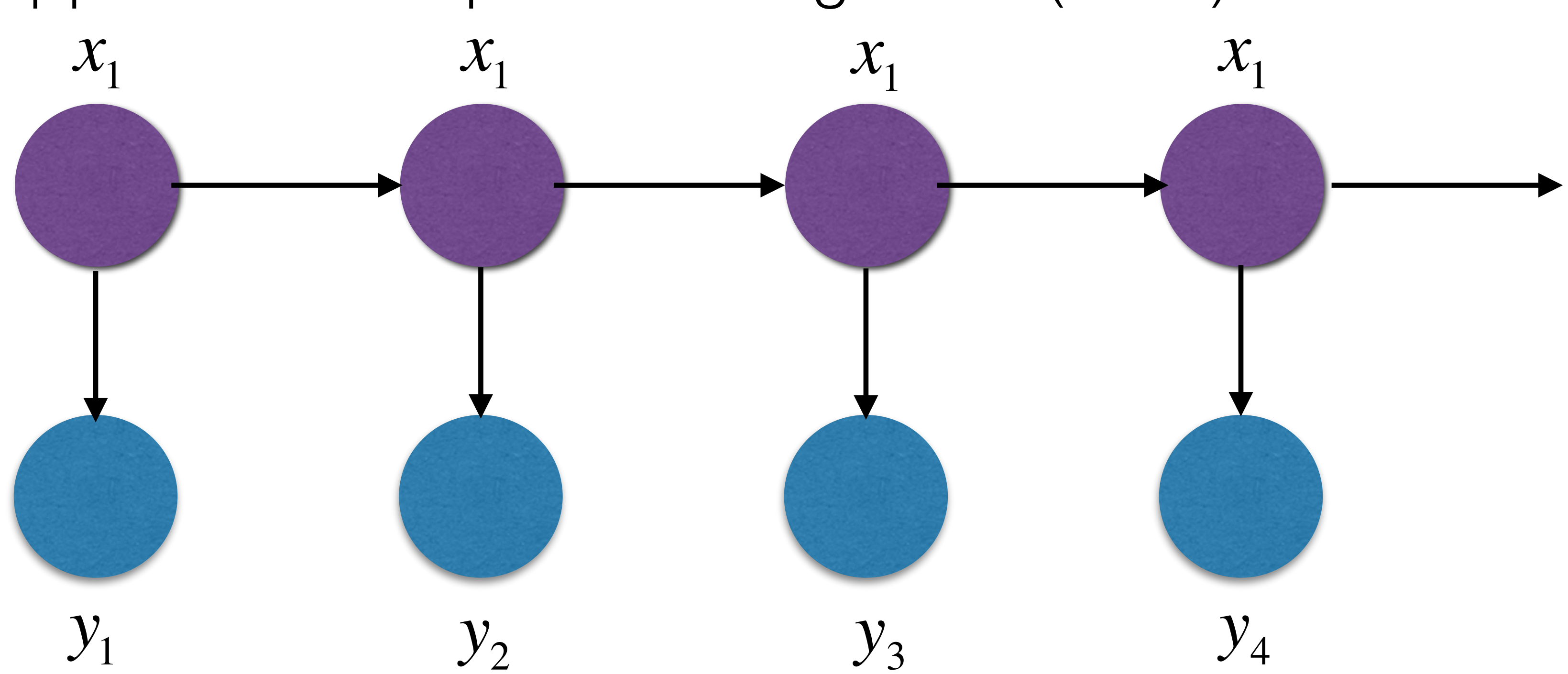
1. **Inference** of the hidden state $p(x_{1:t} \mid y_{1:t}, \theta)$
2. **Learning** of model parameters θ

EM algorithm for HMM

The EM algorithm for HMM is known as the Baum-Welch algorithm

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

- **E-step:** compute the posterior $p(x_{1:t} \mid y_{1:t}, \theta)$ using the forward-backward algorithm
  - Forward pass: recursively computes $p(x_1, x_2, \ldots, x_T \mid y_1, y_2, \ldots, y_T, \theta)$
  - Backward pass: recursively computes $p(x_t, \ldots, x_T \mid y_{t+1}, \ldots, y_T, \theta)$
- **M-step:** estimate all parameters for a fixed distribution of hidden variables.
- For details, see e.g. L. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition” (1989).



HMMs: limitations and extensions

- The basic HMM constructed has only one single hidden state with K different values. This may not be expressive enough for some cases.
- HMM can be extended by allowing a vector of discrete state variables - this is called a factorial HMM.
- A general version of a factorial HMM might have an exponentially large number of parameters. Both complexity and the number of parameters can be further constrained, leading in particular to Dynamic Bayesian Networks.

Hidden Markov Models in Finance

- Regime-change models for portfolio risk: the “regime” is a hidden variable
- Markov-switching models for trend-following: the hidden state is binary: the “bull market” and the “bear market”. See e.g. M. Dai et. al. “Trend Following Trading under a Regime Switching Model” (2010).
- Credit portfolio models with a common unobserved “frailty” factor

Control question

Select all correct answers

1. A Hidden Markov model has a discrete hidden state and a continuous or discrete observed state.
2. A Hidden Markov model has a discrete observed state and a continuous or discrete hidden state.
3. For an EM method with a HMM model, the E-step is called the Forward step, and the M-step is called the Backward step.
4. The Backward step of the EM algorithm for a HMM model returns the model parameters to their values in the previous iteration, to prevent too much variation in the results.

Correct answers: 1