MSML610: Advanced Machine Learning

Stochastic Processes

Instructor: GP Saggese, PhD - gsaggese@umd.edu

References:

Stochastic processes

- **Stochastic processes**

Random Variables and Index Sets

- A stochastic process is a collection of random variables indexed by time or space
- Indexed set: $\{X_t : t \in T\}$ where T can be discrete or continuous
- Describes evolving random phenomena
- Example: daily temperature, stock prices

Markov Chains

- Memoryless stochastic process: $P(X_{t+1}|X_t, ..., X_0) = P(X_{t+1}|X_t)$
- Characterized by a transition matrix P
- Types: discrete-time, continuous-time
- Applications: web ranking, genetic modeling, board games

Stationarity

- A process is stationary if its statistical properties do not change over time
- Strict stationarity: joint distribution invariant under time shift
- Weak stationarity: constant mean and autocovariance depend only on lag
- Important in time-series modeling and signal processing

Martingales

- Process with conditional expectation: $\mathbb{E}[X_{t+1}|X_1, ..., X_t] = X_t$
- Models fair games and conservative estimates
- Used in financial modeling and online learning

Poisson Process

- Models arrival of random events in continuous time
- Events occur independently with constant average rate λ
- Interarrival times are exponentially distributed
- Applications: queuing theory, rare event modeling

Brownian Motion (Wiener Process)

- Continuous-time stochastic process with stationary, independent Gaussian increments
- Starts at 0: $B_0 = 0$
- Used in modeling diffusion, stock prices, reinforcement learning
- Foundation for stochastic differential equations

Autoregressive (AR) Processes

- X_t depends linearly on its past values: $X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t$
- ϵ_t is white noise
- Common in time-series forecasting (e.g., ARIMA)

Moving Average (MA) Processes

- X_t is a linear function of current and past noise: $X_t = \sum_{i=0}^{q} \theta_i \epsilon_{t-i}$
- Used to model short-term dependencies
- Often combined with AR in ARMA/ARIMA models

Hidden Markov Models (HMMs)

- Markov chain with unobserved (hidden) states
- Observations are generated from state-dependent distributions
- Used in speech recognition, bioinformatics, and NLP

Gaussian Processes

- A collection of random variables, any finite number of which are jointly Gaussian
- Fully specified by a mean function and covariance kernel
- Used in Bayesian regression, spatial modeling, and active learning

Stochastic Differential Equations (SDEs)

- Differential equations with random noise
- General form: $dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t$
- Models continuous-time random phenomena
- Applications: physics, finance, control theory

Ergodicity

- Time averages equal ensemble averages under certain conditions
- Ensures statistical estimation is feasible from a single realization
- Crucial for learning from time-series data

Renewal Processes

- Generalization of Poisson processes
- Interarrival times are i.i.d. but not necessarily exponential
- Used in reliability theory and system maintenance

Birth-Death Processes

- Special case of Markov chains with transitions to neighboring states
- Models population dynamics, queue lengths
- Characterized by birth rate λ_n and death rate μ_n

Queueing Models

- Systems where entities wait in line for service
- Described using stochastic processes (e.g., M/M/1 queue)
- Analyzed via arrival and service rate distributions
- Applied in networks, servers, and traffic modeling

Random Walks

- Discrete stochastic process: $X_{t+1} = X_t + \epsilon_t$
- ϵ_t is typically i.i.d. and symmetric
- Central to modeling cumulative processes and diffusion

Time Series Analysis

- Study of data indexed by time with inherent stochasticity
- Techniques: decomposition, smoothing, forecasting
- Used in econometrics, forecasting, and anomaly detection

Law of Large Numbers and Central Limit Theorem

- Justify convergence and Gaussian approximations of stochastic processes
- Law of large numbers: sample mean converges to expected value
- CLT: sum of i.i.d. variables approximates a normal distribution

Monte Carlo Methods

- Use stochastic sampling to approximate expectations and distributions
- Key for Bayesian inference and simulation
- Methods include importance sampling and MCMC

Filtering and Prediction

- Estimate current or future states from noisy observations
- Includes Kalman filters (linear-Gaussian) and particle filters (nonlinear)
- Essential in control, robotics, and state estimation