Johns Hopkins Engineering

625.464 Computational Statistics

Orthogonal Polynomials

Module 10 Lecture 10D

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Orthogonal Polynomials

Can be developed from $1, x, x^2, x^3, \ldots$
by applying G-S w/ approp. weight + normalize $\sqrt{\langle q_i, q_i \rangle}$

Ex/ Legendre Poly Rang $[-1,1]$ $w(x) = 1$

$\tilde{q}_0 = q_0 = 1$ $\langle q_i, q_j \rangle = \int_{-1}^{1} q_i(x) q_j(x)\,dx$

$\tilde{q}_1 = x - \frac{\langle 1, x \rangle}{\langle 1, 1 \rangle} 1$ $\langle 1, x \rangle = \int_{-1}^{1} x\,dx = \left.\frac{x^2}{2}\right|_{-1}^{1} = 0$

$= x$ orthogonal $\circ$ $\langle 1, 1 \rangle = \int_{-1}^{1} 1\,dx = x\Big|_{-1}^{1} = 2$

$\tilde{q}_2 = x^2 - \frac{\langle 1, x^2 \rangle}{\langle 1, 1 \rangle} 1 - \cancel{\frac{\langle x, x^2 \rangle}{\langle x, x \rangle}} x$ $\langle 1, x^2 \rangle = \langle x, x \rangle = \int_{-1}^{1} x^2\,dx = \frac{2}{3}$

$= x^2 - 1/3$ $\langle x, x^2 \rangle = \int_{-1}^{1} x^3\,dx = 0$

Orthogonal Polynomial Recurrences

For the kth polynomial in an orthog set

$\exists$ an a_k (not dep on x) st.

$$q_k(x) - a_k x\, q_{k-1}(x) \text{ is a poly of degree } k-1$$

$$\therefore\ q_k(x) - a_k x q_{k-1} = \sum_{i=0}^{k-1} c_i q_i(x) \quad \text{for some } c_i$$

$$q_k(x) = (a_k x + c_{k-1}) q_{k-1}(x) + c_{k-2} q_{k-2}(x) + \cdots$$

$$c_0 = c_1 = \cdots\ c_{k-3} = 0$$

for some a_k, b_k, c_k

$$\therefore\ q_k(x) = (a_k x + b_k) q_{k-1}(x) - c_k q_{k-2}(x)$$

Polynomial Recurrence Example

Ex/ Legendre Polynomials

$$q_k(x) = \frac{2k-1}{k} x\, q_{k-1}(x) - \frac{k-1}{k} q_{k-2}(x)$$

Using Orthognal Polynomial for Statistics

Q: How to apply O.P. to a data set $X = \{x_1, x_2, \dots, x_n\}$ from an unknown density p?

A: Let's suppose $f(x) = g(x)p(x)$ where $p(x)$ is a prob. density function. Then for any orthonormal set $\{q_i(x)\}$ we can approx f

$$f(x) = \sum_{k=0}^{\infty} c_k q_k(x) \quad \{c_k\}?$$

Finding c_k from **X**

$$f(x) = \sum_{k=0}^{\infty} c_k \, q_k(x)$$

$$c_k = \langle f, q_k \rangle = \int_D q_k(x) g(x) p(x) \, dx$$

$$= E\left[q_k(X) g(X) \right]$$

where $X \sim p$

$\therefore$ $X_1, \dots, X_n$ then we can unbiasedly est c_k by

$$\hat{c}_k = \frac{1}{n} \sum_{i=1}^{n} q_k(x_i) g(x_i)$$

Our Function Estimator... Finally

$$\hat{f}(x) = \frac{1}{n}\sum_{k=0}^{i}\sum_{i=1}^{n} g_k(x_i)\, g(x_i)\, g_k(x)$$

$$f(x) = g(x)\,p(x)$$

iid $x_1, \dots, x_n \sim f$

if $f(x)$ – is a density

$$C_K = E[g_k(X)]$$

$$\hat{f}(x) = \sum_{k=0}^{i} E[g_k(X)]\, g_k(x)$$

$$= \frac{1}{n}\sum_{k=0}^{i}\left[\sum_{i=0}^{n} g_k(x_i)\right] g_k(x)$$