PS5841

Data Science in Finance & Insurance

Smoothing

Yubo Wang

Autumn 2021

Columbia University

Simple Moving Average

$$\hat{s}_t = \frac{1}{m}\left(y_t + y_{t-1} + \cdots + y_{t-(m-1)}\right)$$

COLUMBIA UNIVERSITY

Centered SMA, KNN

$$\hat{s}_t = \frac{1}{m}\left(y_{t+m/2} + \ldots + \ y_t + \cdots + y_{t-m/2}\right)$$

COLUMBIA UNIVERSITY

Exponential Moving Average

$$\hat{s}_t = \frac{y_t + \omega y_{t-1} + \omega^2 y_{t-2} \ldots + \omega^t y_0}{1 + \omega + \omega^2 + \cdots + \omega^t}$$

Since $\lim_{t\to\infty} \sum_{i=0}^{t} \omega^i = \frac{1}{1-\omega}$

$$\hat{s}_t \approx \frac{y_t + \omega y_{t-1} + \omega^2 y_{t-2} \ldots + \omega^t y_0}{1/(1-\omega)}$$

$$\hat{s}_t = \omega \hat{s}_{t-1} + (1-\omega) y_t$$

Columbia University

Whittaker-Henderson Graduation in 1D with order n

- Loss function, balances data fit and smoothness

$$R = F + \lambda S$$

$$= \sum_{i=1}^{N} w_x[r(x) - s(x)]^2 + \lambda \sum_{x=1}^{N} [\Delta^n s(x)]^2$$

$$= (\boldsymbol{r} - \boldsymbol{s})^T \boldsymbol{W} (\boldsymbol{r} - \boldsymbol{s}) + \lambda \boldsymbol{s}^T (\boldsymbol{K}^n)^T (\boldsymbol{K}^n) \boldsymbol{s}$$

COLUMBIA UNIVERSITY

Whittaker-Henderson Graduation in 1D with order n

- Loss function, balances data fit and smoothness

$$R = (\boldsymbol{r} - \boldsymbol{s})^T \boldsymbol{W} (\boldsymbol{r} - \boldsymbol{s}) + \lambda \boldsymbol{s}^T (\boldsymbol{K}^n)^T (\boldsymbol{K}^n) \boldsymbol{s}$$

where

- $\boldsymbol{W}$ is an $N \times N$ diagonal matrix of weights, preferably normalized
- $\boldsymbol{K}^n \boldsymbol{s} = \Delta^n \boldsymbol{s}$ where $\boldsymbol{K}^n$ is a $(N - n) \times N$ matrix

Columbia University

Whittaker-Henderson Graduation in 1D with order n

- Example $N = 6$ and $n = 3$

$$\boldsymbol{K}^3\boldsymbol{s} = \begin{bmatrix} -1 & 3 & -3 & 1 & 0 & 0 \\ 0 & -1 & 3 & -3 & 1 & 0 \\ 0 & 0 & -1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} s(1) \\ s(2) \\ s(3) \\ s(4) \\ s(5) \\ s(6) \end{bmatrix}$$

$$= \begin{bmatrix} \Delta^3 s(4) \\ \Delta^3 s(5) \\ \Delta^3 s(6) \end{bmatrix}$$

Columbia University

Whittaker-Henderson Graduation in 1D with order n

- To minimize the loss,

$$R = (\boldsymbol{r} - \boldsymbol{s})^T \boldsymbol{W} (\boldsymbol{r} - \boldsymbol{s}) + \lambda \boldsymbol{s}^T (\boldsymbol{K}^n)^T (\boldsymbol{K}^n) \boldsymbol{s}$$

$$\frac{\partial R}{\partial \boldsymbol{s}} = -2\boldsymbol{W}(\boldsymbol{r} - \boldsymbol{s}) + 2\lambda (\boldsymbol{K}^n)^T (\boldsymbol{K}^n) \boldsymbol{s} = \boldsymbol{0}$$

$$\hat{\boldsymbol{s}} = [\boldsymbol{W} + \lambda (\boldsymbol{K}^n)^T (\boldsymbol{K}^n)]^{-1} \boldsymbol{W} \boldsymbol{r}$$

Columbia University

Cubic Spline

- $n + 1$ observations (x_i, y_i), $x_0 < x_1 < \cdots < x_n$
- Smoothed function $s(x)$
  - $s(x)$ is twice continuously differentiable
  - $s(x)$ is a 3rd order polynomial on $[x_{i-1}, x_i], i = 1, \ldots n$
  - $s(x_i) = y_i, \forall\ i = 0,1, \ldots, n$
  - $$s(x) = \begin{cases} f_1(x), & x_0 \le x \le x_1 \\ & \cdots \\ f_i(x), & x_{i-1} \le x \le x_i \\ & \cdots \\ f_n(x), & x_{n-1} < x < x_n \end{cases}$$
  - $f_i(x) = a_i + b_i x + c_i x^2 + d_i x^3,\ i = 1, \ldots, n,\ \ d_i \neq 0$

COLUMBIA UNIVERSITY

Cubic Spline

- $4n - 2$ boundary conditions

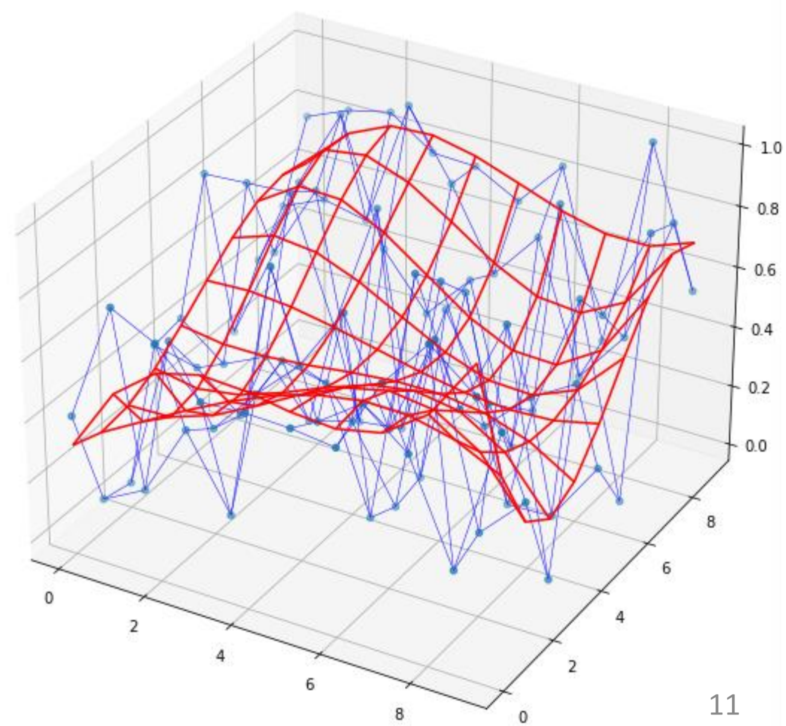
$$f_i(x_{i-1}) = y_{i-1}, \qquad i = 1, \dots n$$
$$f_i(x_i) = y_i, \qquad i = 1, \dots n$$
$$f_i'(x_i) = f_{i+1}'(x_i), \qquad i = 1, \dots n-1$$
$$f_i''(x_i) = f_{i+1}''(x_i), \qquad i = 1, \dots n-1$$

- 2 additional boundary conditions
  - Clamped: assume $f_1'(x_0)$ and $f_n'(x_n)$ are known
  - assume $f_1''(x_0)$ and $f_n''(x_n)$ are known
    - Natural/Simple: $f_1''(x_0) = f_n''(x_n) = 0$
  - Periodic:

$$f_1(x_0) = f_n(x_n), f_1'(x_0) = f_n'(x_n), f_1''(x_0) = f_n''(x_n)$$

Columbia University

B-Spline



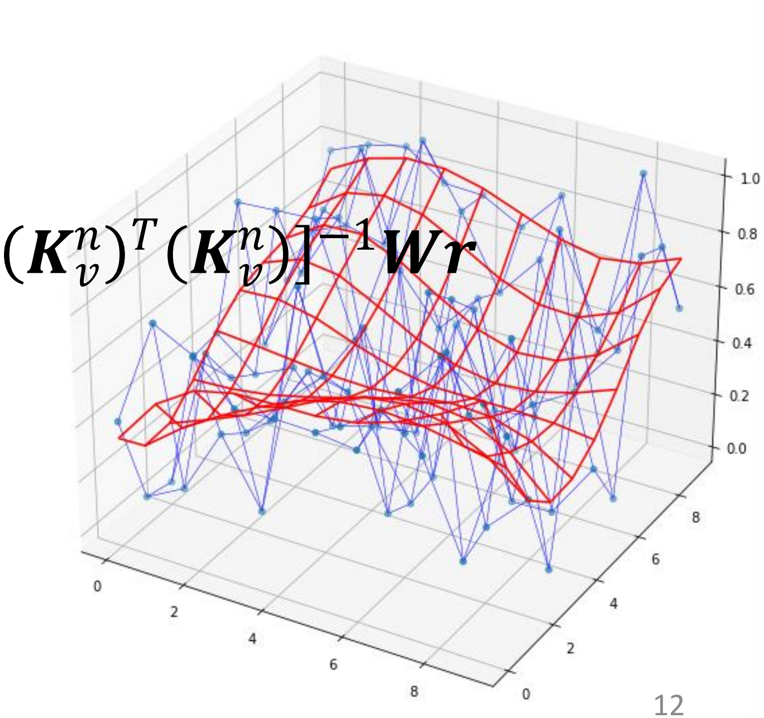
Whittaker-Henderson Graduation in 2D with order (m, n)

- Loss function, balances data fit and smoothness

$$R = F + hS_h + vS_v$$
$$= (\boldsymbol{r} - \boldsymbol{s})^T \boldsymbol{W} (\boldsymbol{r} - \boldsymbol{s}) + h\boldsymbol{s}^T (\boldsymbol{K}_h^m)^T (\boldsymbol{K}_h^m) \boldsymbol{s} + v\boldsymbol{s}^T (\boldsymbol{K}_v^n)^T (\boldsymbol{K}_v^n) \boldsymbol{s}$$

- $\hat{\boldsymbol{s}}$ minimizes the loss,

$$\hat{\boldsymbol{s}} = [\boldsymbol{W} + h(\boldsymbol{K}_h^m)^T (\boldsymbol{K}_h^m) + v(\boldsymbol{K}_v^n)^T (\boldsymbol{K}_v^n)]^{-1} \boldsymbol{W} \boldsymbol{r}$$



COLUMBIA UNIVERSITY

That was