Fast Fourier transforms for efficient computation of kernel estimators

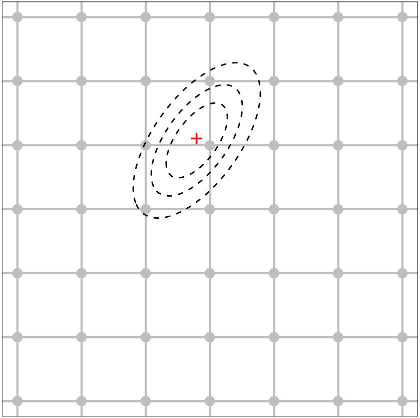
Tarn Duong

LIPN, University of Paris 13, Paris, France

13 April 2016

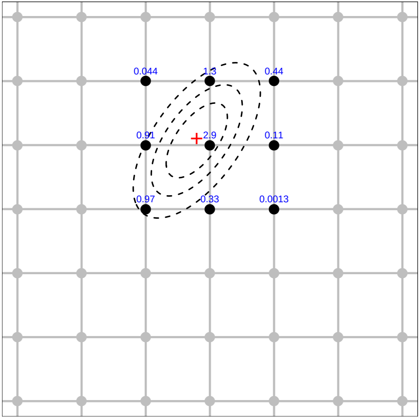
Exact kernel estimation

- Data point $\boldsymbol{X}_1 = (0.266, 0.479)$ (red cross)
- Estimation grid 7×7 on $[0, 1] \times [0, 1]$ (grey circles/lines)
- Gaussian kernel K centred on $\boldsymbol{X}_1$ with variance $\mathbf{H} = [0.01; 0.075; 0.0075, 0.015]$ $K_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{X}_1)$



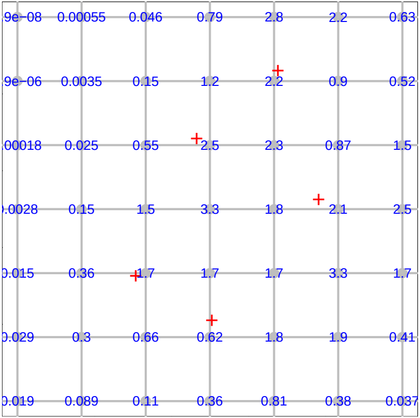
Exact kernel estimation

- Data point $\boldsymbol{X}_1 = (0.266, 0.479)$ (red cross)
- Estimation grid 7×7 on $[0, 1] \times [0, 1]$ (grey circles/lines)
- Gaussian kernel K centred on $\boldsymbol{X}_1$ with variance $\mathbf{H} = [0.01; 0.075; 0.0075, 0.015]$ $K_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{X}_1)$
- Evaluate K (red values) at all grid points that fall in (effective) support (black circles) of $K_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{X}_1)$



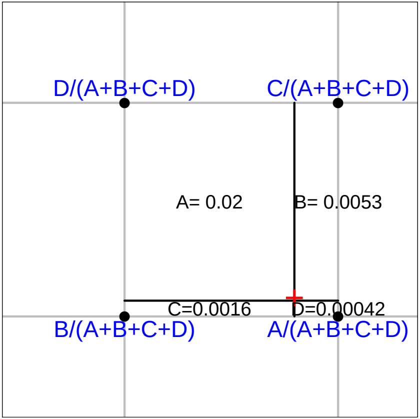
Exact kernel estimation

- Data point $\boldsymbol{X}_1 = (0.266, 0.479)$ (red cross)
- Estimation grid 7×7 on $[0,1] \times [0,1]$ (grey circles/lines)
- Gaussian kernel K centred on $\boldsymbol{X}_1$ with variance $\mathbf{H} = [0.01; 0.075; 0.0075, 0.015]$ $K_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{X}_1)$
- Evaluate K (red values) at all grid points that fall in (effective) support (black circles) of $K_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{X}_1)$
- Repeat for all n data points and sum kernel values $\hat{f}(\boldsymbol{x}; \mathbf{H}) = \frac{1}{n} \sum_{i=1}^{n} K_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{X}_i)$



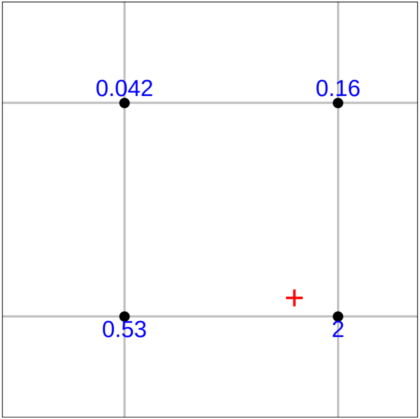
Linear binning

- Data point $X_1 = (0.266, 0.479)$ (red cross)
- Estimation grid 7×7 on $[0,1] \times [0,1]$ (grey circles/lines)
- Zoom in on bin containing X_1



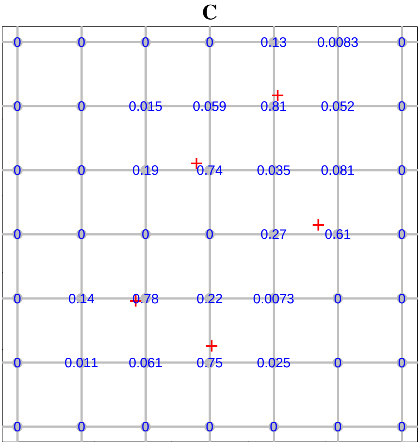
Linear binning

- Data point $X_1 = (0.266, 0.479)$ (red cross)
- Estimation grid 7×7 on $[0, 1] \times [0, 1]$ (grey circles/lines)
- Zoom in on bin containing X_1
- Linear binning converts data point to 2^2 weights (blue values) at bin vertices (black circles)



Linear binning

- Data point $X_1 = (0.266, 0.479)$ (red cross)
- Estimation grid 7×7 on $[0,1] \times [0,1]$ (grey circles/lines)
- Zoom in on bin containing X_1
- Linear binning converts data point to 2^2 weights (blue values) at bin vertices (black circles)
- Repeat for all data points to create $\mathbf{C}$ by summing any overlapping weights



Fast Fourier transform (FFT)

- Kernel estimation = smooth convolution with Gaussian kernel
- Replace smooth convolution with faster discrete convolution (FFT)
- Recall that original grid is indexed by $\ell_1 = 1, \ldots, M_1$ and $\ell_2 = 1, \ldots, M_2$
- Let $\boldsymbol{\delta}$ = dimensions of each bin and $\kappa_{\boldsymbol{\ell}} = \frac{1}{n} K_{\mathbf{H}}(\delta_1 \ell_1, \delta_2 \ell_2), \ell_1, \ell_2 \in \mathbb{Z}$ is the value of Gaussian kernel on 'extended' grid

Fast Fourier transform (FFT)

- Kernel estimation = smooth convolution with Gaussian kernel
- Replace smooth convolution with faster discrete convolution (FFT)
- Recall that original grid is indexed by $\ell_1 = 1, \ldots, M_1$ and $\ell_2 = 1, \ldots, M_2$
- Let $\boldsymbol{\delta}$ = dimensions of each bin and $\kappa_{\boldsymbol{\ell}} = \frac{1}{n} K_{\mathbf{H}}(\delta_1 \ell_1, \delta_2 \ell_2), \ell_1, \ell_2 \in \mathbb{Z}$ is the value of Gaussian kernel on 'extended' grid
- Create zero-padded $P_1 \times P_2$ matrices, e.g. P_1 is smallest power of 2 greater than M_1 etc.

$$\mathbf{K}_0 = \begin{bmatrix} \kappa_{-M_1,-M_2} & \cdots & \kappa_{-M_1,0} & \cdots & \kappa_{-M_1,M_2} & \cdots & \mathbf{0}^T_{P_1-2M_1+1} \\ \vdots & & & \vdots & & & \vdots \\ \kappa_{0,-M_2} & \cdots & \kappa_{0,0} & \cdots & \kappa_{0,M_2} & \cdots & \mathbf{0}^T_{P_1-2M_1+1} \\ \vdots & & & \vdots & & & \vdots \\ \kappa_{M_1,-M_2} & \cdots & \kappa_{M_1,0} & \cdots & \kappa_{M_1,M_2} & \cdots & \mathbf{0}^T_{P_1-2M_1+1} \\ \mathbf{0}_{P_2-2M_2+1} & \cdots & \mathbf{0}_{P_2-2M_2+1} & \cdots & \mathbf{0}_{P_2-2M_2+1} & \cdots & \mathbf{0}_{P_2-2M_2+1,P_2-2M_1+1} \end{bmatrix}$$

$$\mathbf{C}_0 = \begin{bmatrix} \mathbf{0}_{M_1,M_2} & \mathbf{0}_{M_1,M_2} & \mathbf{0}_{M_1,P_2-2M_2} \\ \mathbf{0}_{M_1,M_2} & \mathbf{C} & \mathbf{0}_{M_1,P_2-2M_2} \\ \mathbf{0}_{P_1-2M_1,M_2} & \mathbf{0}_{P_1-2M_1,M_2} & \mathbf{0}_{P_1-2M_1,P_2-2M_2} \end{bmatrix}$$

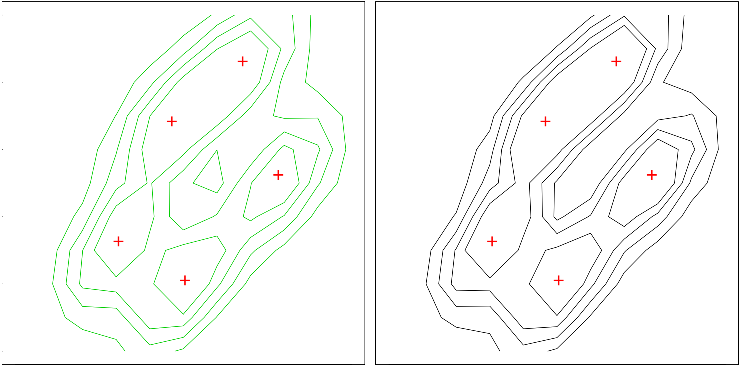
Binned kernel density estimation

- $\mathbf{S} = FFT^{-1}(FFT(\mathbf{C}_0)FFT(\mathbf{K}_0))$ is $P_1 \times P_2$ matrix
- Binned kernel estimator on $M_1 \times M_2$ grid is central $M_1 \times M_2$ elements of $\mathbf{S}$: $\frac{1}{P_1P_2}\mathbf{S}_{(2M_1-1):(3M_1-2),(2M_2-1):(3M_2-2)}$
- (Gramacki & Gramacki, 2016) `arxiv.org/abs/1508.02766`

Binned kernel density estimation

- $\mathbf{S} = FFT^{-1}(FFT(\mathbf{C}_0)FFT(\mathbf{K}_0))$ is $P_1 \times P_2$ matrix
- Binned kernel estimator on $M_1 \times M_2$ grid is central $M_1 \times M_2$ elements of $\mathbf{S}$:
$$\frac{1}{P_1 P_2}\mathbf{S}_{(2M_1-1):(3M_1-2),(2M_2-1):(3M_2-2)}$$
- (Gramacki & Gramacki, 2016) arxiv.org/abs/1508.02766

Binned (7×7 grid) Exact (7×7 grid)



Binned kernel density estimation

- $\mathbf{S} = FFT^{-1}(FFT(\mathbf{C}_0)FFT(\mathbf{K}_0))$ is $P_1 \times P_2$ matrix
- Binned kernel estimator on $M_1 \times M_2$ grid is central $M_1 \times M_2$ elements of $\mathbf{S}$:
$\frac{1}{P_1 P_2}\mathbf{S}_{(2M_1-1):(3M_1-2),(2M_2-1):(3M_2-2)}$
- (Gramacki & Gramacki, 2016) `arxiv.org/abs/1508.02766`