Quantile Regression

Ruibin Xi

Penalized univariate smoothing method

- Assume that

$$y = f(x) + \epsilon$$

- In case of the mean regression, we may consider estimate the function f bv minimizing

$$\text{RSS}(f, \lambda) = \sum_{i=1}^{N} \{y_i - f(x_i)\}^2 + \lambda \int \{f''(t)\}^2 dt,$$

$\lambda = 0:$ f can be any function that interpolates the data.

$\lambda = \infty:$ the simple least squares line fit, since no second derivative can be tolerated.

Penalty method for bivariate smoothing (1)

- Similar to the univariate case, we may consider

$$\sum_{i=1}^{n}(z_i - g(x_i, y_i))^2 + \lambda J\left(g, \Omega, \|\cdot\|_2^2\right),$$

$$J(g, \Omega, \|\cdot\|_2^2) = \int\int_{\Omega} \|\nabla^2 g\|_2^2 dxdy = \int\int_{\Omega} (g_{xx}^2 + 2g_{xy}^2 + g_{yy}^2)dxdy.$$

- The solution to this objective function is called the thin plate smoothing splines

Penalty method for bivariate smoothing (2)

- In general, we may consider minimizing

$$\|\mathbf{y} - \mathbf{g}\|^2 + \lambda J_{md}(g)$$

$\mathbf{y}$ is the vector of y_i data

$\mathbf{g} = (g(\mathbf{x}_1), g(\mathbf{x}_2), \ldots, g(\mathbf{x}_n))'$

$\mathbf{x}$ is a d-vector

$$J_{md} = \int \cdots \int_{\Re^d} \sum_{\nu_1 + \ldots + \nu_d = m} \frac{m!}{\nu_1! \ldots \nu_d!} \left(\frac{\partial^m g}{\partial x_1^{\nu_1} \ldots \partial x_d^{\nu_d}} \right)^2 \mathrm{d}x_1 \ldots \mathrm{d}x_d$$

Penalty method for bivariate smoothing (3)

- It can be shown that (Wahba 200) that the solution to the above problem is of the form

$$g(\mathbf{x}) = \sum_{i=1}^{n} \delta_i \eta_{md}(\|\mathbf{x} - \mathbf{x}_i\|) + \sum_{j=1}^{M} \alpha_j \phi_j(\mathbf{x})$$

$$M = \binom{m+d-1}{d}$$

ϕ_i are linearly independent polynomials of degree less than m

$$\eta_{md}(r) = \begin{cases} \dfrac{(-1)^{m+1+d/2}}{2^{2m-1}\pi^{d/2}(m-1)!\,(m-d/2)!} r^{2m-d} \log(r) & d \text{ even,} \\ \dfrac{\Gamma(d/2-m)}{2^{2m}\pi^{d/2}(m-1)!} r^{2m-d} & d \text{ odd.} \end{cases}$$

Penalty method for bivariate smoothing (4)

- In univariate case, we have considered the penalty to be the total variation penalty

$$P(g) = V(g')$$

where V(f) is the variation of the function f.

- How to extend this to bivariate case?

Penalty method for bivariate smoothing (5)

- For bivariate case, we may consider

$$J(g, \Omega, \|\cdot\|) = V(\nabla g, \Omega, \|\cdot\|) = \int\int_{\Omega} \|\nabla^2 g\| dxdy. \qquad (7.13)$$

- The norm is required to be orthogonal invariance, i.e. $\|U^\top H U\| = \|H\|$ for any symmetric matrix H and orthogonal matrix U (e.g. Frobenius norm)

$$J(g, \Omega, \|\cdot\|_2) = \int\int_{\Omega} \sqrt{g_{xx}^2 + 2g_{xy}^2 + g_{yy}^2} dxdy.$$

Penalty method for bivariate smoothing (6)

Let Ω be a convex, compact region of the plane.

Let Ω be a convex, compact region of the plane, and let Δ denote a collection of sets $\{\delta_i : i = 1, \ldots, N\}$ with disjoint interiors such that $\Omega = \cup_{\delta \in \Delta} \delta$.

When the $\delta \in \Delta$ are planar triangles, Δ is called a triangulation

The continuous functions g on Ω that are linear when restricted to $\delta \in \Delta$ are called triograms

Penalty method for bivariate smoothing (7)

Theorem 7.1. *Suppose that $g : \Omega \to \mathbb{R}$ is a piecewise linear function on the triangulation Δ. For any orthogonally invariant penalty of the form (7.13), there is a constant c dependent only on the choice of the norm such that*

$$J(g, \Omega, \|\cdot\|) = c \sum_e \|\nabla g_e^+ - \nabla g_e^-\| \, \|e\|, \tag{7.15}$$

where e runs over all the interior edges of the triangulation, $\|e\|$ is the Euclidean length of the edge e, and $\|\nabla g_e^+ - \nabla g_e^-\|$ is the Euclidean length of the difference between gradients of g on the triangles adjacent to e.

Penalty method for bivariate smoothing (8)

- The problem

$$\min_{g \in \mathcal{G}_\Delta} \sum |z_i - g(x_i, y_i)| + \lambda J_\Delta(g)$$

can be formulated as

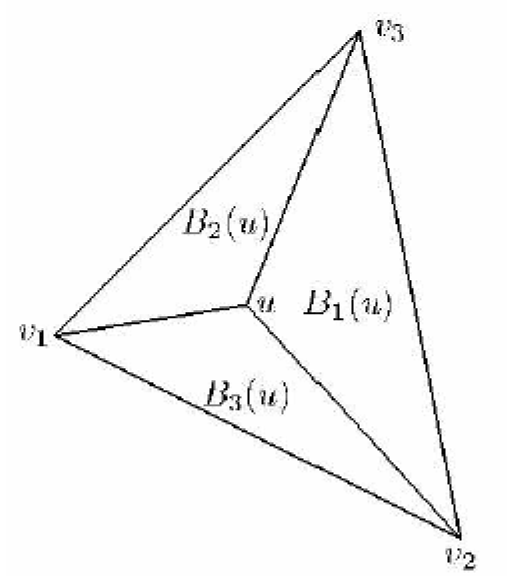
$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \left| z_i - a_i^\top \beta \right| + \lambda \sum_{k=1}^{M} \left| h_k^\top \beta \right| .$$

Penalty method for bivariate smoothing (9)

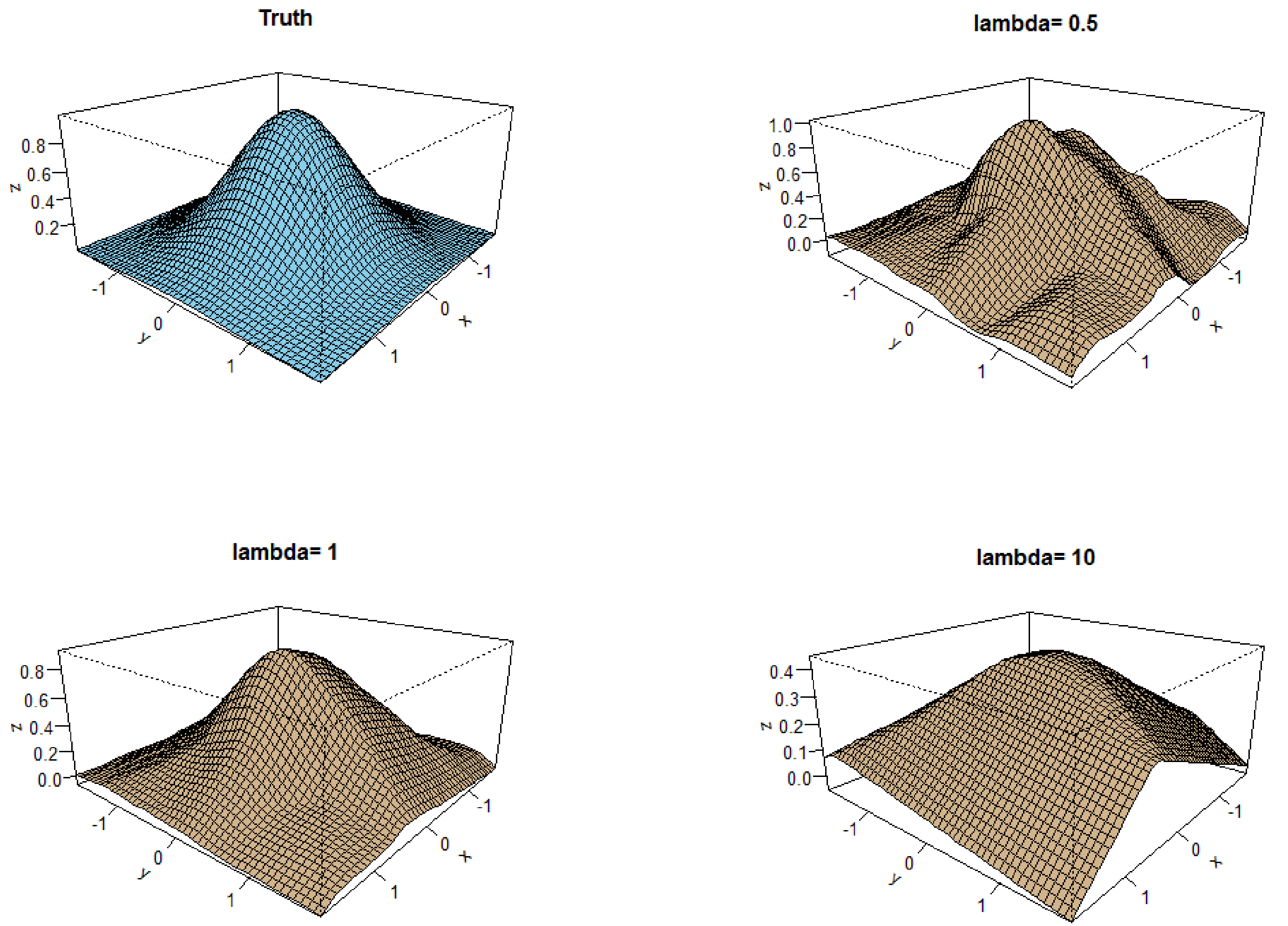
- Barycentric coordinates

$$u_j = \sum_{i=1}^{3} B_i(u) v_{ij} \quad j = 1, 2, \qquad B_1(u) = \frac{A(u, v_2, v_3)}{A(v_1, v_2, v_3)}$$

$$A(v_1, v_2, v_3) = \frac{1}{2} \begin{vmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ 1 & 1 & 1 \end{vmatrix}$$

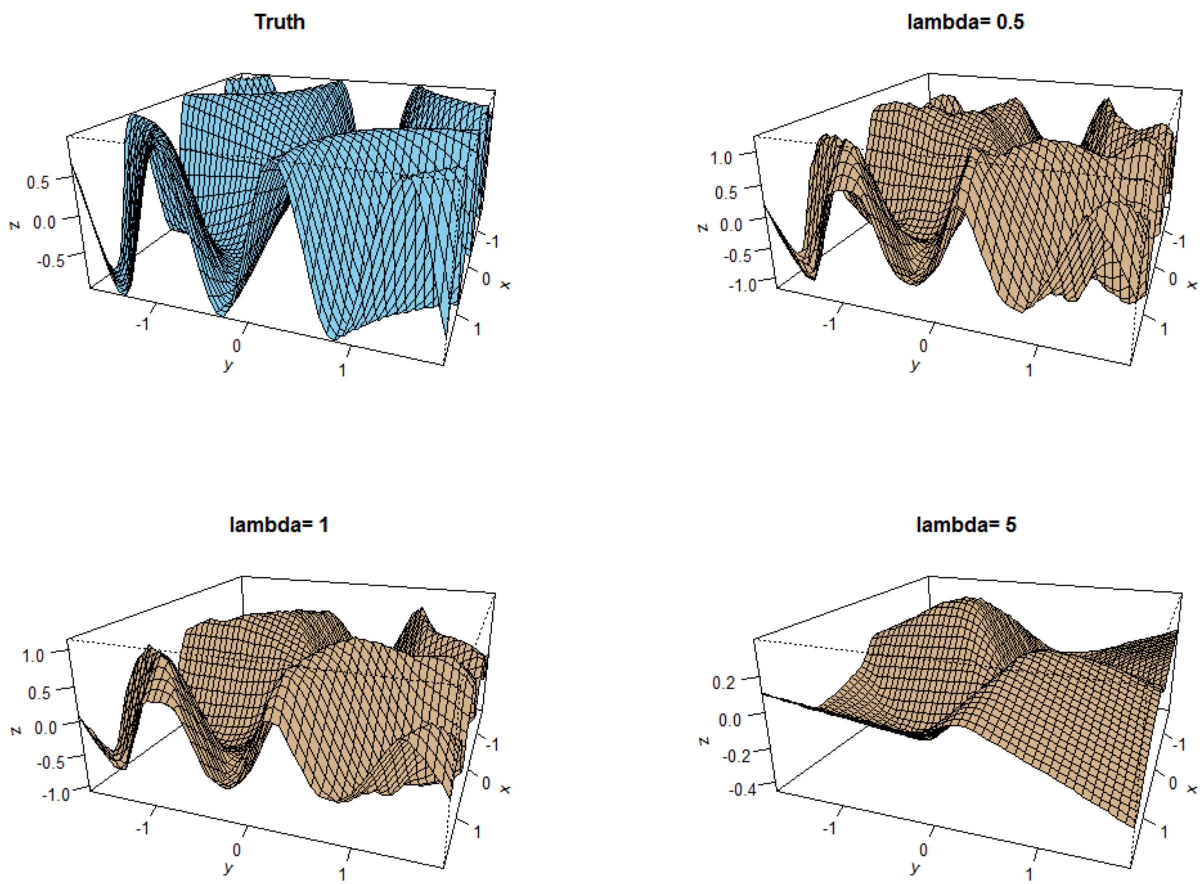


Examples (1)



f(x,y) = exp(-x^2-y^2)
z = f(x,y) + error

Example (2)



f(x,y) =sin(pi*x*y)
z = f(x,y) + error

Examples (3)