Robust counterparts of errors-in-variables problems

G Alistair Watson

Division of Mathematics

University of Dundee

Given (1) Data $(y_i, \mathbf{x}_i), i = 1, \ldots, m$

$$y_i \in R, \mathbf{x}_i \in R^t$$

(2) Model $y = \sum_{j=1}^{n} a_j \phi_j(\mathbf{x})$.

Find $\mathbf{a} \in R^n$.

Assumption: Errors in all variables.

Total least squares

$A : A_{ij} = \phi_j(\mathbf{x}_i), i = 1, \ldots, m, j = 1, \ldots, n$

$\mathbf{y} + \mathbf{r} = (A + E)\mathbf{a}$ (model equations)

minimize $\|E : \mathbf{r}\|_F^2$.

Robust counterpart

Uncertainty set $\mathcal{E}$, $(\mathbf{y} + \mathbf{r}, A + E) \in \mathcal{E}$.

$$\min_{\mathbf{a}} \max_{(\mathbf{y}+\mathbf{r}, A+E) \in \mathcal{E}} \|\mathbf{y} + \mathbf{r} - (A + E)\mathbf{a}\|$$

Interpretation: minimize $\|\tilde{\mathbf{y}} - \tilde{A}\mathbf{a}\|$ w.r.t. $\mathbf{a}$ over the **worst** of perturbations defined by $(\tilde{\mathbf{y}}, \tilde{A}) \in \mathcal{E}$.

$\|.\|$ denotes l_2 or Frobenius norm.

Example

$$\mathcal{E} = \{\mathbf{y} + \mathbf{r}, A + E : \|\mathbf{r}\| \le \rho_1, \|E\| \le \rho_2\}.$$

$$\min_{\mathbf{a}} \max_{\|\mathbf{r}\| \le \rho_1, \|E\| \le \rho_2} \|\mathbf{y} + \mathbf{r} - (A + E)\mathbf{a}\|.$$

Solution $\mathbf{a}$ minimizes

$$\|\mathbf{y} - A\mathbf{a}\| + \rho_2 \|\mathbf{a}\|.$$

Good methods available: Golub, El Ghaoui,...

TLS : $\mathbf{y} \to \mathbf{y} + \mathbf{r}$, $A \to A + E$,

$$\phi_j(\mathbf{x}_i) \to \phi_j(\mathbf{x}_i) + E_{ij},$$

$$\mathbf{y} + \mathbf{r} = (A + E)\mathbf{a}.$$

(Structure on E: Beck, van Huffel,..)

Errors-in-variables: $\mathbf{y} \to \mathbf{y} + \mathbf{r}$,

$$\phi_j(\mathbf{x}_i) \to \phi_j(\mathbf{x}_i + \mathbf{s}_i).$$

$$y_i + r_i = \sum_{j=1}^{n} a_j \phi_j(\mathbf{x}_i + \mathbf{s}_i), i = 1, \ldots, m.$$

(Model equations)

Minimise $\sum_{i=1}^{m} r_i^2 + \sum_{i=1}^{m} \|\mathbf{s}_i\|^2$.

Orthogonal distance regression (ODR)

Robust Counterpart

Uncertainty set $\mathcal{E}$.

$(\mathbf{y}+\mathbf{r}, \mathbf{x}_i+\mathbf{s}_i, i=1,\ldots,m) \in \mathcal{E}$.

Problem: find

$$\min_{\mathbf{a}} \max_{(\mathbf{y}+\mathbf{r}, \mathbf{x}_i+\mathbf{s}_i, i=1,\ldots,m)\in\mathcal{E}} \|\mathbf{v}\|,$$

$$v_i = y_i + r_i - \sum_{j=1}^{n} a_j \phi_j(\mathbf{x}_i + \mathbf{s}_i), i = 1, \ldots, m.$$

Interpretation: if

$$\tilde{z}_i = \tilde{y}_i - \sum_{j=1}^{n} a_j \phi_j(\tilde{\mathbf{x}}_i), i = 1, \ldots, m,$$

minimize $\|\tilde{\mathbf{z}}\|$ w.r.t. $\mathbf{a}$ over the **worst** of perturbations defined by $(\tilde{\mathbf{y}}, \tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_m) \in \mathcal{E}$.

Ideal. Nonlinear in $\mathbf{s}_i$. Assume $\phi_j \in C^1$. Then for all i,

$$v_i = w_i + O(\|\mathbf{s}_i\|^2),$$

where

$$\begin{aligned} w_i &= y_i + r_i - \sum_{j=1}^{n} a_j \phi_j(\mathbf{x}_i) - \sum_{j=1}^{n} a_j \nabla_{\mathbf{x}} \phi_j(\mathbf{x}_i) \mathbf{s}_i \\ &= z_i + r_i - \sum_{j=1}^{n} a_j \nabla_{\mathbf{x}} \phi_j(\mathbf{x}_i) \mathbf{s}_i, \\ z_i &= y_i - \sum_{j=1}^{n} a_j \phi_j(\mathbf{x}_i). \\ \mathbf{z} &= \mathbf{y} - A\mathbf{a}. \end{aligned}$$

Problem Find

$$\min_{\mathbf{a}} \max_{(\mathbf{y}+\mathbf{r}, \mathbf{x}_i+\mathbf{s}_i, i=1,\ldots,m) \in \mathcal{E}} \|\mathbf{w}\| \qquad (1)$$

for different uncertainty sets.

Pointwise uncertainty

Define $\mathbf{s}^T = (\mathbf{s}_1^T, \ldots, \mathbf{s}_m^T)$. Let

$$S_1 = \{(\mathbf{r}, \mathbf{s}) : |r_i| \le \rho_i, |(\mathbf{s}_i)_j| \le \gamma_{ij},$$

$$i = 1, \ldots, m, j = 1, \ldots, t\},$$

where $\rho_i, \gamma_{ij}, i = 1, \ldots, m, j = 1, \ldots, t$ are given.

Uncertainty set

$$\mathcal{E}_1 = \{\mathbf{y} + \mathbf{r}, \mathbf{x}_i + \mathbf{s}_i, i = 1, \ldots, m : (\mathbf{r}, \mathbf{s}) \in S_1\}.$$

Problem:

$$\min_{\mathbf{a}} \max_{(\mathbf{r},\mathbf{s}) \in S_1} \|\mathbf{w}\|. \tag{2}$$

Define

$$G_i = \begin{bmatrix} \nabla_{\mathbf{x}} \phi_1(\mathbf{x}_i) \\ \nabla_{\mathbf{x}} \phi_2(\mathbf{x}_i) \\ .. \\ \nabla_{\mathbf{x}} \phi_n(\mathbf{x}_i) \end{bmatrix} \in R^{n \times t}, i = 1, \ldots, m,$$

and let

$$D_i = \text{diag}\{\gamma_{i1}, \ldots, \gamma_{it}\}, i = 1, \ldots, m.$$

Theorem 1 Let $\mathbf{a}^* \in R^n$ solve

$$\text{minimize } \|\mathbf{c}\|, \text{ where} \qquad (3)$$

$$c_i = |z_i| + \rho_i + \|D_i G_i^T \mathbf{a}\|_1, i = 1, \ldots, m.$$

Then $\mathbf{a}^*$ solves (2).

Proof Involves upper bound on each $|w_i|$, which we can show can be attained. ■

Tractable? Problem is:

$$\begin{array}{rcl} \text{minimize} & h & \text{subject to} \\ \|\mathbf{u}\| & \leq & h, \\ |z_i| + \rho_i + \|D_i G_i^T \mathbf{a}\|_1 & \leq & u_i, i = 1, \ldots, m. \end{array}$$

Second order cone programming problem (SOCP). Good interior point methods exist (Boyd,..)

l_1 norm in (2). Linear l_1 problem:

$$\min_{\mathbf{a}} \{\|\mathbf{y} - A\mathbf{a}\|_1 + \|M\mathbf{a}\|_1\},$$

(M depends on data).

Chebyshev norm in (2). SOCP.

Now define

$$S_2 = \{(\mathbf{r}, \mathbf{s}) : |r_i| \le \rho_i, \|\mathbf{s}_i\| \le \gamma_i, i = 1, \ldots, m\},$$

where $\rho_i, \gamma_i, i = 1, \ldots, m$ are given.

(Same as before if $t = 1$).

$$\mathcal{E}_2 = \{\mathbf{y} + \mathbf{r}, \mathbf{x}_i + \mathbf{s}_i, 1 = 1, \ldots, m : (\mathbf{r}, \mathbf{s}) \in S_2\},$$

$$\min_{\mathbf{a}} \max_{(\mathbf{r},\mathbf{s}) \in S_2} \|\mathbf{w}\|. \tag{4}$$

Theorem 2 Let $\mathbf{a}^* \in R^n$ be a solution to the problem

$$\text{minimize } \|\mathbf{c}\|, \text{ where} \tag{5}$$

$$c_i = |z_i| + \rho_i + \gamma_i \|G_i^T \mathbf{a}\|, i = 1, \ldots, m.$$

Then $\mathbf{a}^*$ solves (4).

Proof Much as before. ∎

Problem can be restated:

minimize h subject to

$$\begin{aligned} \|\mathbf{t}\| &\leq h, \\ \gamma_i \|G_i^T \mathbf{a}\| &\leq t_i - z_i - \rho_i, \ i = 1, \ldots, m, \\ \gamma_i \|G_i^T \mathbf{a}\| &\leq t_i + z_i - \rho_i, \ i = 1, \ldots, m. \end{aligned}$$

SOCP.

Can replace bounds on $\mathbf{s}_i$ by $\|\mathbf{s}_i\|_A \leq \gamma_i$ for arbitrary norm; can use other norms on $\mathbf{w}$. Eg l_∞, l_1 gives linear l_1 problem; l_2, l_∞, SOCP.

Normwise uncertainty

Define

$$S_3 = \{(\mathbf{r}, \mathbf{s}) : \|\mathbf{r}\| \leq \rho, \|\mathbf{s}_i\| \leq \gamma_i, i = 1, \ldots, m\},$$

$$\mathcal{E}_3 = \{\mathbf{y} + \mathbf{r}, \mathbf{x}_i + \mathbf{s}_i, i = 1, \ldots, m : (\mathbf{r}, \mathbf{s}) \in S_3\},$$

$$\min_{\mathbf{a}} \max_{(\mathbf{r},\mathbf{s}) \in S_3} \|\mathbf{w}\|. \qquad (6)$$

For any $\mathbf{a}$ and $(\mathbf{r}, \mathbf{s}) \in S_3$,

$$\mathbf{w} = \mathbf{z} + \mathbf{r} - \sum_{i=1}^{m} \mathbf{e}_i \mathbf{a}^T G_i \mathbf{s}_i,$$

$$\|\mathbf{w}\| \leq \|\mathbf{z}\| + \rho + \sum_{i=1}^{m} \gamma_i \|G_i^T \mathbf{a}\|.$$

Upper bound sum of Euclidean norms (minimized by interior point methods). But not attained.

Natural question Since TLS case requires we minimize

$$\|\mathbf{y} - A\mathbf{a}\| + \rho_2 \|\mathbf{a}\|.$$

What does this form achieve in the e-i-v case?

For any $\mathbf{s}$, define M by

$$M^T = [G_1 \mathbf{s}_1, \ldots, G_m \mathbf{s}_m].$$

Theorem 3. Let $\mathbf{a}^*$ minimize

$$\|\mathbf{y} - A\mathbf{a}\| + \gamma\|\mathbf{a}\|. \tag{7}$$

Then if $G_i^T\mathbf{a}^* \neq 0, i = 1, \ldots, m$, $\mathbf{a}^*$ solves the problem

$$\min_{\mathbf{a}} \max_{\|\mathbf{r}\| \leq \rho, \frac{\|M\mathbf{a}\|}{\|\mathbf{a}\|} \leq \gamma} \|\mathbf{w}\|. \tag{8}$$

Proof. Not difficult. ∎

Artificial uncertainty set.

Define

$$\mathbf{d}^T = [r_1, \ldots, r_m, \mathbf{s}_1^T, \ldots, \mathbf{s}_m^T],$$

so that $\mathbf{d} \in R^{m(1+t)}$.

$$\mathcal{E}_4 = \{\mathbf{y} + \mathbf{r}, \mathbf{x} + \mathbf{s}, \|\mathbf{d}\| \leq \rho\}.$$

$$\min_{\mathbf{a}} \max_{\|\mathbf{d}\| \leq \rho} \|\mathbf{w}\|. \tag{9}$$

Now let

$$E(\mathbf{s}) = \begin{bmatrix} \mathbf{s}_1^T G_1^T \\ .. \\ \mathbf{s}_m^T G_m^T \end{bmatrix}.$$

Then

$$\begin{aligned} \mathbf{w} &= \mathbf{y} - A\mathbf{a} + \mathbf{r} - E(\mathbf{s})\mathbf{a} \\ &= \mathbf{y} + \mathbf{r} - (A + E(\mathbf{s}))\mathbf{a} \\ &= \mathbf{b}(\mathbf{d}) - G(\mathbf{d})\mathbf{a} \end{aligned}$$

where

$$G(\mathbf{d}) = A + \sum_i d_i A_i,$$

$$\mathbf{b}(\mathbf{d}) = \mathbf{y} + \sum_i d_i \mathbf{b}_i.$$

Problem is to find

$$\min_{\mathbf{a}} \max_{\|\mathbf{d}\| \leq \rho} \|G(\mathbf{d})\mathbf{a} - \mathbf{b}(\mathbf{d})\|. \qquad (10)$$

Structured robust problem (El Ghaoui,..). Pose as semi-definite programming problem (interior point methods).

Note: l_2 norm on perturbation size implies a correlated bound.

Variants are:

$$\min_{\mathbf{a}} \max_{\|\mathbf{d}\|_\infty \le \rho} \|G(\mathbf{d})\mathbf{a} - \mathbf{b}(\mathbf{d})\|_\infty,$$

or

$$\min_{\mathbf{a}} \max_{\|\mathbf{d}\|_1 \le \rho} \|G(\mathbf{d})\mathbf{a} - \mathbf{b}(\mathbf{d})\|_1.$$

These can be posed as LP problems (Hindi and Boyd).

Nonlinear models

$$y = f(\mathbf{a}, \mathbf{x}).$$

Problem: find

$$\min_{\mathbf{a}} \max_{(\mathbf{y}+\mathbf{r}, \mathbf{x}_i+\mathbf{s}_i, i=1,\ldots,m) \in \mathcal{E}} \|\mathbf{v}\|,$$

$$v_i = y_i + r_i - f(\mathbf{a}, \mathbf{x}_i + \mathbf{s}_i), i = 1, \ldots, m.$$

As before, settle for solving

$$\min_{\mathbf{a}} \max_{(\mathbf{y}+\mathbf{r}, \mathbf{x}_i+\mathbf{s}_i, i=1,\ldots,m) \in \mathcal{E}} \|\mathbf{w}\| \qquad (11)$$

where

$$w_i = y_i + r_i - f(\mathbf{a}, \mathbf{x}_i) - \nabla_{\mathbf{x}} f(\mathbf{a}, \mathbf{x}_i)\mathbf{s}_i, i = 1, \ldots, m.$$

Particular choice of $\mathcal{E}$ gives problem

$$\min_{\mathbf{a}} \{\|\mathbf{y} - \mathbf{f}(\mathbf{a}, \mathbf{x})\| + \rho\|\mathbf{a}\|\}$$

Gauss-Newton method. Subproblem:

$$\min_{\mathbf{d}} \{\|\mathbf{y} - A\mathbf{d}\| + \rho\|\mathbf{d}\|\}.$$

Familiar problem from TLS case.

Concluding Comments

- Extends robust counterparts of TLS problems to e-i-v context (deal explicitly with perturbations of variable values)

- Range of uncertainty sets

- The problem (1) can be solved for range of these

- Pose as standard convex problems (SOCP, SDP) for which good IP methods are available

- Other norms can be dealt with (eg pose as LP)

- Can extend to nonlinear problems