Stat 771 Spring 09, UW Madison, MN.

Notes on Nelder-Mead algorithm for <u>maximizing</u> a log likelihood $\ell(\theta)$, with parameter $\theta \in \mathbb{R}^p$

The algorithm produces a sequence of sets $S_1, S_2 \dots S_m \dots$ with any $S = \{\theta_1, \dots \theta_{p+1}\}$, $p+1$ points in $\mathbb{R}^p$

Assume that we start with S not a hyperplane.

Update until stopping condition:

1. θ_i's sufficiently close within S
2. $\ell(\theta_i)$'s " " " "
3. reach maximum number of iterations

} <u>or</u>

Updates proceed by trying to improve the $\ell(\theta)$ value of points in current $S_M = S$

<u>Reflect step</u> First identify three points in S

$\theta_s =$ value with smallest $\ell(\theta)$

$\theta_b =$ " " best (largest) $\ell(\theta)$

$\theta_{s+1} =$ " " 2^{nd} smallest $\ell(\theta)$

and form $\theta_0 = \frac{1}{p} \sum_{\theta_i \neq \theta_s} \theta_i$ [mean of non-worst]

Construct $\quad \theta_r = (1+\alpha)\,\theta_0 - \alpha\,\theta_\ell$

for some $\alpha > 0$

$$\theta_0 = \frac{1}{P} \sum_{\theta_j \neq \theta_\ell} \theta_j$$

Expansion Step:

If $\quad \ell(\theta_r) \geq \ell(\theta_b)$ [reflection ~~expansion~~ improves things]

expand further to

$$\theta_e = \theta_0 + \gamma(\theta_r - \theta_0)$$

for some $\gamma > 1$

if $\ell(\theta_e) > \ell(\theta_r)$

Return

$$\mathcal{S}_{m+1} = \{\mathcal{S}_m \mid \theta_\ell\} + \{\theta_e\}$$

[i.e drop θ_ℓ and include θ_e]

if $\ell(\theta_e) \leq \ell(\theta_r)$

Return

$$\mathcal{S}_{m+1} = \{\mathcal{S}_m \mid \theta_\ell\} + \{\theta_r\}$$

i.e., if reflection improves things, try further improvement and take the better of θ_r, θ_e.

Intermediate Case

[reflection not too bad]

if

$$l(\theta_{n+1}) \leq l(\theta_r) < l(\theta_b)$$

then set

$$\mathcal{S}_{m+1} = \{\mathcal{S}_m \setminus \theta_n\} + \{\theta_r\}$$

Contraction cases

1. External $l(\theta_n) \leq l(\theta_r) < l(\theta_{n+1})$

form

$$\theta_c = \theta_0 + \beta(\underbrace{\theta_r - \theta_0}_{\theta_r - \theta_0})$$

for some $\beta \in (0,1)$

if $l(\theta_c) > l(\theta_n)$

$$\mathcal{S}_{m+1} = \{\mathcal{S}_m \setminus \theta_n\} + \{\theta_c\}$$

else "shrink" [see below]

2. Internal $l(\theta_r) < l(\theta_n)$ [reflected point very bad]

form

$$\theta_{cc} = \theta_0 + \beta(\theta_n - \theta_0)$$

if $\ell(\theta_{cc}) > \ell(\theta_s)$

$$S_{m+1} = \{S_m \setminus \theta_s\} + \{\theta_{cc}\}.$$

else "shrink"

Shrink step. if no luck above, and directed here,

$$S_{m+1} = \{ \theta_b , \quad \forall j = b, \quad \theta_b + \delta(\theta_j - \theta_b) \}$$

for some $\delta \in (0,1)$

Note: in constrained parameter spaces, some adjustment is required to keep all points $\theta_r, \theta_0, \theta_e, \theta_c, \theta_{cc}$. in the parameter space

---