Variable Selection For Discrete Competing Risks Models

THE CHOICE OF PENALTY TERM

Lupeng Kong
2015180132290 76

- The form of parameters for $\lambda_{\mathrm{r}}(t|x)$
  - $\eta_{itr} = \beta_{0tr} + x_i^T \gamma_r, t = 1, \ldots q; r = 1, \ldots, m,$
    - $x_i^T = (x_{i1}, \ldots, x_{ip})$
    - $\gamma_r^T = (\gamma_{r1}, \ldots, \gamma_{rp})$
- The penalized ML estimation:
  - $l_{\zeta_1,\zeta_2}(\beta_0, \gamma) = l(\beta_0, \gamma) - J_{\zeta_1,\zeta_2}(\beta_0, \gamma)$
  - $l(\beta_0, \gamma)$ is the original ML estimation
  - Aim for variable selection

- The original penalty term

- $$J_{\zeta_1,\zeta_2}(\beta_0,\gamma) = \zeta_1 \sum_{r=1}^{m} \sum_{t=2}^{q} (\beta_{0tr} - \beta_{0,t-1,r})^2 + \zeta_2 \sum_{j=1}^{p} \phi_j ||\gamma_{.j}||$$
  $$= \zeta_1 J_1(\beta_0) + \zeta_2 J_2(\gamma)$$

- $\zeta_1 J_1(\beta_0)$: smooth the cause−specific baseline coefficients β_0

- $\zeta_2 J_2(\gamma)$: variable selection (lasso)

- $||\gamma_{.j}|| = \sqrt{\gamma_{1j}^2 + \gamma_{2j}^2 + \ldots + \gamma_{mj}^2} = 0$ ➡ x_j will be removed

- Improvement for penalty item $J_1(\beta_0)$
  - To reduce the number of time periods q in $J_1(\beta_0)$
  - expanded in low−rank B−spline basis function
    - $\beta_{0tr} = \sum_{s=1}^{d_r} \alpha_{0sr} B_s(t)$
  - The improved penalty term 1:
    - $J_{\zeta_1,\zeta_2}(\alpha_0, \gamma) = \zeta_1 \sum_{r=1}^{m} \sum_{s=2}^{d_r} \left(\alpha_{0sr} - \alpha_{0,s-1,r}\right)^2 + \zeta_2 \sum_{j=1}^{p} \phi_j ||\gamma_{.j}||$
    - $d_r <$ q

- Improvement for penalty item $\zeta_2 J_2(\gamma)$
  - The corresponding penalization is small for strong predictors and large for weak predictors.
  - penalized estimators with adaptive weights can provide consistent variable selection.
  - The final penalty term:
    - $J_{\zeta_1,\zeta_2}(\alpha_0,\gamma) = \zeta_1 \sum_{r=1}^{m} \sum_{s=2}^{d_r} (\alpha_{0,s,r} - \alpha_{0,s-1,r})^2 + \zeta_2 \sum_{j=1}^{p} \phi_j^a ||\gamma_{.j}||$
    - $\phi_j^a = \frac{\sqrt{m}}{||\gamma_{.j}^{Init}||}$,
  - $\left||\gamma_{.j}^{Init}|\right|$: penalized estimator that results form penalty function above with $\zeta_2 = 0$.