Advanced Survival Modeling Approaches for Prognostic Analysis via Multi-Omic Data Integration

C. Angelini[a], D. De Canditiis[a], I. De Feis[a] and Antonella Iuliano[b]

[a] *Consiglio Nazionale delle Ricerche, Istituto per le Applicazioni del Calcolo "M. Picone", Italy*
[b] *University of Basilicata, Departmant of Health Science, Italy*

High-throughput omics technologies have significantly advanced biomedical research by enabling comprehensive molecular profiling. While single omics layers provide limited information, their integration offers a multidimensional view of disease biology, improving prognostic modeling and supporting personalized medicine, particularly in oncology.
Survival analysis based on the Accelerated Failure Time (AFT) model is a powerful framework for linking time-to-event outcomes with molecular predictors. Integrating multi-omics data within this model allows the investigation of how molecular alterations across different data views jointly affect patient survival and recurrence, facilitating biomarker discovery and improving survival prediction.
In this talk, we propose a novel cooperative survival analysis approach, termed CoopAFT, which extends the AFT model through a penalized likelihood framework for joint estimation and variable selection. The log-linear AFT model is defined as

$$\log T_i = \boldsymbol{u}_i^T \boldsymbol{\beta_u} + \boldsymbol{z}_i^T \boldsymbol{\beta_z} + \sigma\varepsilon_i, \quad i = 1, \dots, n,$$

where $\varepsilon_i \in \mathbb{R}$ is a random error independent from the two omics data views $\boldsymbol{u}_i$ and $\boldsymbol{z}_i$, $\boldsymbol{\beta_u}$ and $\boldsymbol{\beta_z}$ are the corresponding regression coefficients and σ is a scale parameter.
Model parameters $\boldsymbol{\theta} = (\boldsymbol{\beta_u}^T, \boldsymbol{\beta_z}^T, \sigma)^T \in \mathbb{R}^p \times \mathbb{R}^+$, with $p = p_u + p_z$, are estimated by minimizing a penalized negative AFT log-likelihood, combining a cooperative penalty that enforces agreement between the two data views and an ℓ_1-penalty promoting sparsity, i.e.,

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\mathbb{R}^{\boldsymbol{p}}\times\mathbb{R}^+}{\operatorname{argmin}} \left\{ -\frac{1}{n}\ell(\boldsymbol{\theta}) + \mathcal{P}_{\lambda,\rho}(\boldsymbol{\theta}) \right\},$$

where the penalty term $\mathcal{P}_{\lambda,\rho}(\cdot)$ is given by

$$\mathcal{P}_{\lambda,\rho}(\boldsymbol{\theta}) = (1-\rho)\lambda\|\mathbf{U}\boldsymbol{\beta_u} - \mathbf{Z}\boldsymbol{\beta_z}\|^2 + \rho\lambda(||\boldsymbol{\beta_u}||_1 + ||\boldsymbol{\beta_z}||_1).$$

Parameter $\lambda > 0$ is the regularization parameter tuned by some data-driven criterion, while $\rho \in [0, 1]$, is a user-defined parameter representing the trade-off between the two penalties. The optimization problem is efficiently solved using a proximal gradient descent algorithm. We establish theoretical consistency of the proposed estimator and assess its performance through simulation studies and real cancer survival data analyses.
Overall, the CoopAFT framework enhances survival prediction by integrating complementary multi-omics information, supporting robust biomarker identification and personalized treatment strategies in cancer research.

References

1. Angelini, C., De Canditiis, D., De Feis, I., and Iuliano, A. (2024). A Network-Constrain Weibull AFT Model for Biomarkers Discovery. *Biometrical Journal*, **66(7)**, e202300272.

2. Ding, D. Y., Li, S., Narasimhan, B., & Tibshirani, R. (2022). Cooperative learning for multiview analysis. *Proceedings of the National Academy of Sciences*, **119(38)**, e2202113119.