

BIOF3 FIGCODE · SCI 绘图代码集

LASSO 交叉验证曲线 (LASSO CV Curve)

导出日期：2026年6月27日



扫码关注微信公众号【生信F3】
获取文章完整内容，分享生物信息学
最新知识。

分类：生存分析 | 依赖包：glmnet、survival

解决的生物学问题

我的 LASSO 模型在什么正则化强度下预测最优？应该选多少个基因进 signature？

应用场景

- LASSO Cox 论文 Figure 2A
- 选择最优 lambda 确定 signature 基因数
- 展示模型复杂度与预测性能的权衡
- 多 cohort 验证时对比 CV 曲线

输入数据格式

cv.glmnet 对象 (family = "cox")

关键参数

- nfolds 折数 (默认 10)
- lambda.min 最小 deviance 对应 lambda
- lambda.1se 一个标准误内最简模型
- alpha = 1 (纯 LASSO)

快速开始

1. 直接在线运行

打开 [FigCode 在线绘图](#)，点击本工具卡片的"在线绘图"按钮即可用内置示例数据出图，零环境配置。

2. 本地复现

下载脚本和示例数据，本地 RStudio 运行：

```
# 下载
curl -O https://<your-site>/figcode/scripts/lasso-cv-plot.R
curl -O https://<your-site>/figcode/data/lasso-cv-plot.csv
```

3. 安装依赖

```
# CRAN 包
install.packages(c("glmnet", "survival"))

# Bioconductor 包 (如需)
# BiocManager::install(c())
```

完整代码

```
library(glmnet)
library(survival)

# --- Demo data ---
set.seed(42)
n <- 200; p <- 50
x <- matrix(rnorm(n * p), n, p)
colnames(x) <- paste0("Gene", 1:p)
time <- rexp(n, 0.05) * exp(-0.3 * x[,1] + 0.2 * x[,2])
event <- rbinom(n, 1, 0.7)
y <- Surv(time, event)

# --- LASSO Cox CV ---
cv_fit <- cv.glmnet(x, y, family = "cox", alpha = 1, nfolds = 10)

# --- Plot ---
plot(cv_fit, main = "LASSO Cox Cross-Validation")
abline(v = log(cv_fit$lambda.min), col = "red", lty = 2)
abline(v = log(cv_fit$lambda.1se), col = "blue", lty = 2)
legend("topright", legend = c("lambda.min", "lambda.1se"),
      col = c("red", "blue"), lty = 2, bty = "n")
```

替换为自己的数据

脚本中以 `# --- Demo data ---` 标注的段落是示例数据生成代码。替换为自己的数据时，**保持列名一致**即可：

- 输入格式：cv.glmnet 对象 (family = "cox")
- 使用 `read.csv()` / `readRDS()` 读取本地文件

延伸阅读

- 相关教程：[LASSO 交叉验证曲线 完整流程](#)
- 出现 bug? 欢迎在 [FigCode 页面](#) 点击对应卡片，在评论区留言。