Source task and conventions

Prove reuse of R for all ten-fold fits without changing the penalized model.

Conventional solution and proof

Compute $X = RV^\top$ once with V an orthonormal basis containing the span of every data row. For a training-index subset T, $X_T = R_T V^\top$. Every coefficient decomposes as $V\theta + \eta$ with $V^\top \eta = 0$, so the fold objective equals

$$\sum_{i \in T} L(y_i, \beta_0 + r_i^\top \theta) + \lambda \|\theta\|^2 + \lambda \|\eta\|^2.$$

For $\lambda > 0$, the minimizer has $\eta = 0$ and is exactly the fit on R_T. Held-out predictions also agree because $X_i V\theta = r_i^\top \theta$. This argument applies independently to every fold and every positive λ, proving reuse. The reduction is an exact orthogonal coordinate change on the complete row span, not a rank truncation. It preserves even each fold's directions that are unobserved in its training rows, since positive regularization makes their fitted coefficient zero.