

RidgeLRGLH: Testing General Linear Hypotheses

Haoran Li

1 Model and hypothesis

The package tests $BC = 0$ in the multivariate model $Y = BX + E$. Responses form a $p \times n$ matrix, the design is $m \times n$, and contrasts form an $m \times q$ matrix. The design and contrast matrices must have full row and column rank, respectively. Include an intercept as a row of the design when appropriate. Residual and tested degrees of freedom are $n_2 = n - m$ and $n_1 = q$.

The test uses the largest eigenvalue of $W_1(W_2 + \lambda I_p)^{-1}$, where both Wishart matrices are normalized by their respective degrees of freedom. Under the conditions in Li (2025), $p^{2/3}(\ell_{\max} - \Theta_1)/\Theta_2$ has an asymptotic real Tracy–Widom distribution. This approximation may be inaccurate for small dimensions, small tested rank, irregular spectral edges or violations of the model.

2 An example

Here the first row of the design is an intercept and the hypothesis jointly tests the remaining ten predictors. Simulated data are generated under the null.

```
> library(RidgeLRGLH)
> set.seed(8)
> Y <- matrix(rnorm(20 * 80), 20)
> X <- rbind(1, matrix(rnorm(10 * 80), 10))
> C <- diag(11)[, -1, drop = FALSE]
> fit <- ridge_gllht(Y, X, C, selection = "minimax",
+                   control = list(size = 60, grid_size = 80))
> fit
```

Ridge-regularized largest root test (minimax)

```
data: Y with hypothesis B C = 0
TW = -0.28058, lambda = 5.0571, p-value = 0.2245
alternative hypothesis: greater
```

```
> head(fit$tuning)
```

	lambda	center	scale	score
1	0.02931676	9.665145	11.287381	0.1132573
2	0.58796381	4.206976	3.723629	0.1872672
3	1.14661087	2.824162	2.364335	0.2079639
4	1.70525792	2.148023	1.757165	0.2175008
5	2.26390498	1.739861	1.406043	0.2228641
6	2.82255203	1.464698	1.175081	0.2262447

For production use, the default spectral grid has 500 masses and 500 transform points. The smaller grids above keep this example fast.

3 Regularization

A fixed ridge can be supplied as `lambda`. Its default is the mean residual eigenvalue. Adaptive choices use a finite candidate grid: `bayes_identity`, `bayes_covariance`, or `minimax`. The default candidates range from $\min(0.1p/n_2, 1)\text{tr}(W_2)/p$ to $5\text{tr}(W_2)/p$ at ten equally spaced points. Minimax optimizes the worst signal-to-noise ratio over nonnegative mixtures of identity and covariance priors normalized to equal total signal strength. Adaptive calibration requires the consistency and separation conditions of Section 4 of the article. It does not give an exact finite-sample guarantee or adjust for testing multiple hypotheses.

4 Estimation and diagnostics

Algorithm 1 estimates a population spectral distribution by linear programming, matching the real and imaginary relative discrepancies of Q_1 and Q_2 . Default grid selection and weight truncation follow Supplement S.2. The scalar parameterization of Section 3 yields the edge and derivatives by bracketed root finding and implicit differentiation, without requiring a dense ODE trajectory. Only `lpSolve` and `RMTstat` are required beyond standard R packages.

```
> fit$spectrum$loss
[1] 0.3727324
> sum(fit$spectrum$weights)
[1] 1
> fit$calibration
```

center	scale	beta	rho	s
0.90163330	0.71466660	2.66390382	5.74505355	0.26312219
s_prime	s_double_prime			
0.07045842	0.03835507			

Inspect sensitivity to spectral grid size, fitting ridge and constraint order. The loss is an approximation diagnostic before weight truncation; it has no universal significance threshold. A user-supplied population spectrum can be used for oracle studies:

```
> ridge_gllht(Y, X, C, lambda = 1,
+             spectrum = list(sigma = 1, weights = 1))
```

Ridge-regularized largest root test (fixed)

```
data: Y with hypothesis B C = 0
TW = -1.3522, lambda = 1, p-value = 0.5266
alternative hypothesis: greater
```

5 Precomputed matrices

Use `ridge_root(W1, W2, n1, n2)` for a common-covariance double Wishart problem. Both matrices must be positive semidefinite and normalized by their respective degrees of freedom. The denominator may be singular, but must have positive trace. This interface does not turn arbitrary sample covariance pairs into an equality-of-covariance test.

6 Reference

Li, H. (2025). Ridge-Regularized Largest Root Test For High-Dimensional General Linear Hypotheses. Accepted by the Annals of Statistics. <https://arxiv.org/abs/2504.15510>.