Optimum Value of h for Regularization of Estimated Covariance Matrix

Hossein Zareamoghaddam

April 05, 2016

Motivation: Consider linear regression models of the form:

$$\mathbf{Y}_{m,n} = \mathbf{X}_{m,n}\,\mathbf{U}_{m,n} + \epsilon_{m,n}, \;\; m = 1, \ldots, M, \;\; n = 1, \ldots, N \tag{1}$$

$$\hat{\Sigma}_P = \frac{1}{M(T_m - K - Q)} \sum_{m=1}^{M} \mathbf{R}_m^T \mathbf{R}_m$$

$$\tilde{\Sigma}_P = h\,\mathrm{diag}(\hat{\Sigma}_P) + (1-h)\,\hat{\Sigma}_P \tag{2}$$

$$\mathbf{B}_m = \hat{\mathbf{U}}_m \tilde{\Sigma}_P^{-1/2}$$

OBJECTIVE:

We aim to obtain more efficient shrinkage parameter h, $(0 \leq h \leq 1)$, which will give the highest split-half reliability for prewhitening.

Data set

The data set (given by Dr Diedrichsen) is of the collection of 12 hemispheres (6 subjects $\times$ 2 hemispheres) for each contralateral and ipsilateral: Let

$k = 318, 375, 319, 351, 326, 286, 328, 263, 310, 349, 277, 296$

$\mathbf{Y}$	$984 \times k$	Time time-series data for each hemisphere
$\mathbf{T}$	80×3	Number of “digit", “hand" and “run"
$\mathbf{X}intercept$	$123 \times 1 \times 6 \times 8$	Intercepts for the runs
$\mathbf{X}task$	$123 \times 10 \times 6 \times 8$	The task-based regressors for Left and Right fingers.
$\mathbf{X}hpf$	$123 \times 5 \times 6 \times 8$	High-pass filter regressors for each run.

- Dissimilarity distance: **pairwise cross-validation Mahalanobis distance**

$$d^2(\mathbf{b}_k, \mathbf{b}_j) = (\mathbf{b}_k - \mathbf{b}_j)_A \Sigma_A^{-1} (\mathbf{b}_k - \mathbf{b}_j)_B^T$$

- (Split half) Reliability of dissimilarity: **Pearson correlation with fix intercept**

$$r_i = \frac{\mathbf{d}_{odd}^T \mathbf{d}_{even}}{\sqrt{\mathbf{d}_{odd}^T \mathbf{d}_{odd}} \sqrt{\mathbf{d}_{even}^T \mathbf{d}_{even}}} \qquad , i = 1, \ldots, 12$$

- Statistical Inferences on r_i's: **Apply Fisher Z-transformation**

$$z = 0.5 log \frac{1+r}{1-r} \qquad , \qquad r = \frac{exp(2z) - 1}{exp(2z) + 1}$$

has normal distribution with SD: $1/\sqrt{N-1} = 1/\sqrt{12-1}$

- If $\bar{z}$ is the average of z_i's, then $\bar{r}$ is the average reliability measure of dissimilarity (likewise C.I. can be obtained)

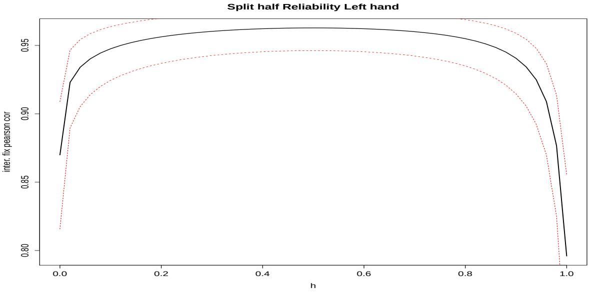


Figure: Split-half reliability (for different *h*-values) of the **Left** hand

Split-half reliabilities for different h-values (Left)

```
    Reliability Low C.I.  Up C.I.    h
1    0.8698777 0.8157496 0.9089021 0.00
2    0.9229937 0.8896570 0.9465437 0.02
3    0.9340727 0.9052959 0.9543150 0.04
...
24   0.9627236 0.9461043 0.9742859 0.46
25   0.9627960 0.9462080 0.9743361 0.48
26   0.9628217 0.9462448 0.9743540 0.50   **
27   0.9628004 0.9462143 0.9743392 0.52
28   0.9627312 0.9461152 0.9742912 0.54
...
48   0.9248635 0.8922908 0.9478572 0.94
49   0.9089084 0.8698865 0.9366240 0.96
50   0.8766121 0.8250233 0.9137101 0.98
51   0.7959557 0.7157522 0.8554357 1.00
```

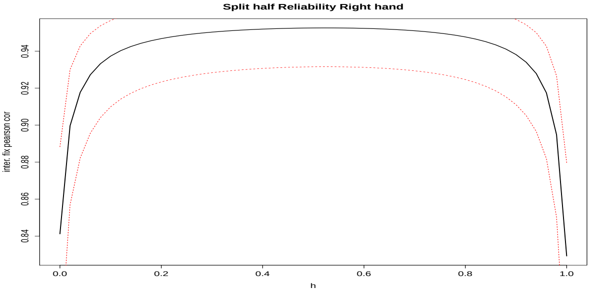


Figure: Split-half reliability (for different *h*-values) of the **Right** hand

```
Split-half reliabilities for different h-values (Right

      Reliability Low C.I.  Up C.I.    h
1     0.8411994 0.7765672 0.8883108 0.00
2     0.8996591 0.8569719 0.9300859 0.02
3     0.9177318 0.8822565 0.9428432 0.04
...
...
24    0.9524149 0.9313602 0.9671212 0.46
25    0.9525045 0.9314880 0.9671835 0.48
26    0.9525614 0.9315692 0.9672231 0.50   **
27    0.9525856 0.9316038 0.9672400 0.52
28    0.9525768 0.9315912 0.9672339 0.54
...
...
49    0.9174681 0.8818861 0.9426576 0.96
50    0.8948247 0.8502430 0.9266608 0.98
51    0.8292359 0.7603684 0.8796646 1.00
```

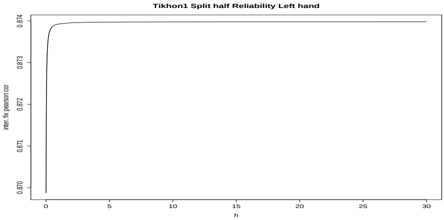
Tikhonov regularization

$$\mathbf{A}\,\mathbf{x} = \mathbf{y} \qquad \Rightarrow \qquad \min_x \{ \parallel \mathbf{A}\,\mathbf{x} - \mathbf{b} \parallel_2^2 + h \parallel \mathbf{L}\,\mathbf{x} \parallel_2^2 \}$$

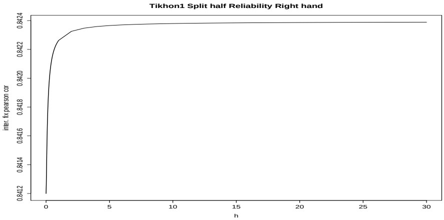
$$\mathbf{x}_{L,h} = \left(\mathbf{A}^T\mathbf{A} + h\mathbf{L}^T\mathbf{L}\right)^{-1}\mathbf{A}^T\mathbf{y}$$

Some options of **L**:

- $\mathbf{L} = \mathcal{I}$
- $$\mathbf{L}_2 = \frac{1}{4}\begin{pmatrix} -1 & 2 & -1 & & & 0 \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ 0 & & & -1 & 2 & -1 \end{pmatrix}$$

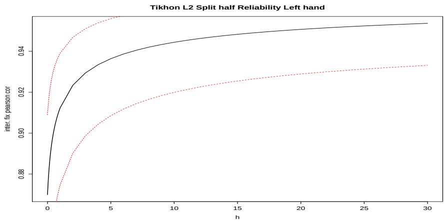


(a) Left

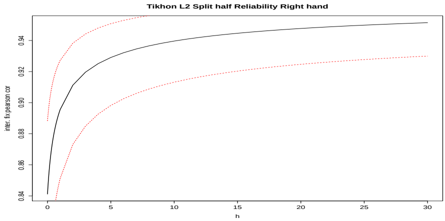


(b) Right

Figure: Reliabilities of Tikhonov Reg. ($\mathbf{L} = \mathcal{I}$)

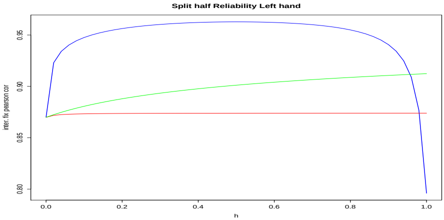


(a) Left

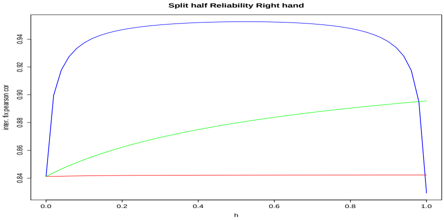


(b) Right

Figure: Reliabilities of Tikhonov Reg. ($\mathbf{L} = \mathbf{L}_2$)



(a) Left



(b) Right

Figure: Reliabilities of Shrinkage and Tikhonov Reg.

Thank you!

- Walther, A., Nilli, H., Ejaz, N., Alink, A., Kriegeskorte, N., Diedrichsen, J. (2016). Reliability of dissimilarity measures for multivariate fMRI pattern analysis, NeuroImage.
- Ledoit, O., Wolf, M., (2004). Honey, i shrunk the sample covariance matrix. J. Portf. Manag. http://dx.doi.org/10.3905/jpm.2004.110.
- Hearn, T.A., Reichel, L. (2014). Application of denoising methods to regularization of ill-posed problems, Numer. Algorithms, 66, pp. 761-777.
- Noschese, S., Reichel, L. (2012). Inverse problems for regularization matrices, Numer. Algorithms, 60 , pp. 531-544.