









|                             |  |
|-----------------------------|--|
|                             | <ul style="list-style-type: none"> <li>• Statistical learning</li> <li>• Challenges concerning high-dimensional models and differences from low-dimensional models</li> <li>• Classical variable selection techniques for linear regression models: <math>R^2</math>, adj.<math>R^2</math>, <math>C_p</math></li> <li>• Information criteria selection: KL divergence, AIC/TIC/BIC derivation</li> <li>• Cross-validation based selection: Leave-one-out and K-fold</li> <li>• Under- and overfitting or the bias-variance trade-off</li> <li>• Ridge shrinkage: theoretical properties, bias/variance trade-off, GCV</li> <li>• Lasso shrinkage: regularization paths, LARS, coordinate descent algorithm, prediction error bounds, degrees of freedom for lasso, support recovery, stability selection, knock-offs; inference by debiasing, post-selection inference, Bayesian inference</li> <li>• Extensions of Lasso: elastic net, group lasso, adaptive lasso, fused lasso</li> <li>• Other techniques: sparse graphical models, sparse PCA, sparse Discriminant Analysis</li> </ul> |
| Inline resources            | <p>Moodle website of the class : LSTAT2450 - Statistical learning. Estimation, selection and inference.<br/> <a href="https://moodleucl.uclouvain.be/course/view.php?id=14890">https://moodleucl.uclouvain.be/course/view.php?id=14890</a></p>   |
| Bibliography                | <ul style="list-style-type: none"> <li>• Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.</li> <li>• James, G., Witten, D., Hastie, T., and Tibshirani, R. (2014). An Introduction to Statistical Learning: With Applications in R. Springer</li> <li>• Hastie, T., Tibshirani, R. and Wainwright, M. J. (2015). Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman and Hall/CRC.</li> <li>• Wainwright, M. J. (2019). High-Dimensional Statistics: A Non-Asymptotic Viewpoint. Cambridge University Press.</li> <li>• Bühlmann, P., van de Geer, S. (2011). Statistics for High-Dimensional Data. Springer.</li> </ul>   |
| Faculty or entity in charge | LSBA   |

| Programmes containing this learning unit (UE)                                      |         |         |              |   |
|--|---------|---------|--------------|---|
| Program title  | Acronym | Credits | Prerequisite | Learning outcomes   |
| Master [120] in Statistics:<br>Biostatistics                                       | BSTA2M  | 5       |              |  |
| Master [120] in Mathematics  | MATH2M  | 5       |              |  |
| Master [120] in Statistics:<br>General   | STAT2M  | 5       |              |  |
| Master [120] in Data Science<br>Engineering  | DATE2M  | 5       |              |  |
| Certificat d'université :<br>Statistique et science des<br>données (15/30 crédits) | STAT2FC | 5       |              |  |
| Master [120] in Data Science:<br>Information Technology                            | DATI2M  | 5       |              |  |