TTK4260 - Introduction to Multivariate Data Modelling
Spring 2025

Big Data Cybernetics Gang

Q & A about the digital exam

Question L1.1

(content units associated to this question: internal validity, external validity)

What do we mean with internal and external validity? May you make some practical examples?

Question L1.2

(content units associated to this question: least squares)

What are least squares estimators? What is the geometrical intuition behind them? And how can they be formulated mathematically?

Question L1.3

(content units associated to this question: maximum likelihood)

What are maximum likelihood estimators? What is the probabilistic intuition behind them? And how can they be formulated mathematically?

Question L1.4

(content units associated to this question: training set, test set, validation set)

Why should one divide a dataset into training, test, and validation sets? And how should one select such sets from the original dataset?

Question L1.5

(content units associated to this question: cross validation)

What does cross validation mean? How should it be used? And why should one use a cross validation approach instead of using a training, test, and validation sets based approach?

Question L1.6

(content units associated to this question: maximum a posteriori)

What are maximum a posteriori estimators? What is the probabilistic intuition behind them? And how can they be formulated mathematically?

Question L1.7

(content units associated to this question: least squares, maximum likelihood, maximum a posteriori)

What are the conditions about existence and uniqueness of the LS, ML and MAP estimates?

Question L1.8

(content units associated to this question: maximum likelihood)

Derive the maximum likelihood estimator for the separable problem $y_i = \theta u_i + v_i$ with $v_i \sim \mathcal{N}\left(0, \sigma^2\right)$ with σ^2 known, θ unknown and deterministic

Question L1.9

(content units associated to this question: bias vs variance tradeoff)

Derive and comment the bias-variance tradeoff

Question L1.10

(content units associated to this question: model order selection)

What do we expect to see when training and testing different estimators with different model order complexities? How should we account for the effects that we see on the statistical performance indexes?

Question L1.11

(content units associated to this question: statistical performance indexes)

Which statistical performance indexes would you consider when dealing with a regression problem? And which peculiarities / usages do they have?

Question L1.12

(content units associated to this question: statistical performance indexes)

Which statistical performance indexes would you consider when dealing with a classification problem? And which peculiarities / usages do they have?

Question L1.13

(content units associated to this question: factorial design)

What does “design of experiments” mean? And “factorial design”? Which alternative factorial-design based alternatives do you know, and what are the tradeoffs among them?

Question L1.14

(content units associated to this question: PCA)

What does PCA mean from a geometrical point of view? How is it formulated mathematically, and how does it connect with SVD?

Question L1.15

(content units associated to this question: PCA)

What are the uses of PCA? And how can its results be interpreted?

Question L1.16

(content units associated to this question: PCA)

What are the uses of the loadings plots and scores plots, in a PCA?

Question L1.17

(content units associated to this question: PCA, model order selection, outliers detection)

How can one decide how many components should be used when analysing some data through PCA? And how can one decide whether a sample is an outlier or not, through PCA?

Question L1.18

(content units associated to this question: model order selection)

How does the Ockham's razor principle connect with the model order selection problem? Which alternative strategies can be used to solve the model order selection problem?

Question L1.19

(content units associated to this question: rotated PCA)

What does "rotated PCA" mean? How does this concept connect with PCA, from both geometrical and mathematical points of view?

Question L1.20

(content units associated to this question: independent component analysis)

Which type of problems does the ICA algorithm solve? Which assumptions does it require? And how does it work, from intuitive perspectives?

Question L1.21

(content units associated to this question: total least squares)

What does "total least squares" mean? How does this concept connect with least squares, from both geometrical and mathematical points of view?

Question L1.22

(content units associated to this question: ANOVA)

What does ANOVA mean? Which type of analyses does this approach serve? How is it formulated mathematically, and how can one use it in practice?

Question L1.23

(content units associated to this question: partial least squares)

What does PLS mean? Which type of analyses does this approach serve? How is it formulated mathematically, and how can one use it in practice?

Question L1.24

(content units associated to this question: multiple linear regression)

What does MLR mean? Which type of analyses does this approach serve? How is it formulated mathematically, and how can one use it in practice?

Question L1.25

(content units associated to this question: principal component regression)

What does PCR mean? Which type of analyses does this approach serve? How is it formulated mathematically, and how can one use it in practice?

Question L1.26

(content units associated to this question: principal component regression, multiple linear regression, partial least squares)

How do PLS, MLR, and PCR relate to each other? In which cases does one expect one of them to work better than the other ones, and viceversa?

Question L1.27

(content units associated to this question: NIPALS)

What is the NIPALS algorithm? How does it work, from a graphical perspective? Which advantages does it bring over SVD, when used to compute a PCA?

Question L1.28

(content units associated to this question: metamodelling)

What does metamodelling mean? When would one want to use a metamodelling approach? What are the potential shortcomings of a metamodel? *(won't be asked for the year 2022)*

Question L1.29

(content units associated to this question: stationarity, ergodicity)

What do "stationarity" and "ergodicity" mean? Why are these two concepts important when dealing with statistical analyses of time series? And what would the lack of stationarity and ergodicity imply in practice?

Question L1.30

(content units associated to this question: linear time invariant systems, discrete time systems)

Which LTI model structures do you know that are suitable to do control-oriented modelling of discrete time MISO systems?

Question L1.31

(content units associated to this question: prediction error methods)

What is the principle behind prediction error methods? When should the focus be on prediction errors, when identifying a dynamical model?

Question L1.32

(content units associated to this question: prediction error methods, model order selection)

What are the implications of choosing an ARX, instead of an ARMAX, instead of an OE model structure when doing system identification? And what are the implications of choosing different model orders? How should one choose a specific structure and order?

Question L1.33

(content units associated to this question: Hammerstein Wiener models)

What are Hammerstein Wiener models, and what are their usages?

Question L1.34

(content units associated to this question: p values)

What is a p value? How should it be computed? What is its usage for? And its drawbacks?

Question L1.35

(content units associated to this question: statistical test)

What is a statistical test? How can it be interpreted from geometrical perspectives? And from mathematical perspectives?

Question L1.36

(content units associated to this question: statistical test, significance of a test, power of a test, uniformly most powerful test)

What are the statistical performance indexes associated to a statistical test? And which concepts may one use to say that a test is "better" than another one?

Question L1.37

(content units associated to this question: simple hypothesis, composite hypothesis)

What are the differences between simple and composite hypotheses? How do the formulations of hypothesis testing algorithms change, depending on which type of hypothesis is considered?

Question L1.38

(content units associated to this question: Linear Discriminant Analysis)

What does the Linear Discriminant Analysis algorithm do? How? Which advantages and disadvantages does it have?

Question L1.39

(content units associated to this question: Partial least squares discriminant analysis)

What does the Partial least squares discriminant analysis algorithm do? How? Which advantages and disadvantages does it have?

Question L1.40

(content units associated to this question: Logistic regression)

What does the logistic regression algorithm do? How? Which advantages and disadvantages does it have?

Question L1.41

(content units associated to this question: Support Vector Classifiers)

What does the Support Vector Classification algorithm do? How? Which advantages and disadvantages does it have?

Question L1.42

(content units associated to this question: kernel trick)

What is the kernel trick? Where may one use it, and why?

Question L1.43

(content units associated to this question: K-means classification)

What does the K-means algorithm do? How? Which advantages and disadvantages does it have?

Question L1.44

(content units associated to this question: DBSCAN classification)

What does the DBSCAN algorithm do? How? Which advantages and disadvantages does it have? What are its differences with the k-means algorithm?

Question L1.45

(content units associated to this question: decision trees, random forests)

What do decision trees and random forests do? How? Which advantages and disadvantages do they have?

Question L1.46

(content units associated to this question: variables selection, features selection)

What is the need for features selection? And what are the differences between feature engineering and selection?

Question L1.47

(content units associated to this question: variables selection, features selection)

What are the pros and cons of using a wrapper method for feature selection, or an embedded method for the same sake?

Question L1.48

(content units associated to this question: variables selection, features selection)

why is categorical cross entropy loss used instead of mean squared error as a cost function in classification problems as against the regression problems? *(won't be asked for the year 2022)*

Question L1.49

(content units associated to this question: regularization)

Which regularization technique is more effective in feature selection? And why?