Skolkovo Institute of Science and Technology, Skoltech
http://www.skoltech.ru/

Course “Experimental data processing”
October, 2018

Course instructor: Tatiana Podladchikova
Contact: t.podladchikova@skoltech.ru

Teaching assistant: Dmitry Shadrin
Contact: dmitry.shadrin@skoltech.ru

The goal of this course: to introduce students to practically useful approaches of data processing for control and forecasting. The focus will be on identifying the hidden and implicit features and regularities of dynamical processes using experimental data.
The course exposes data processing methods from multiple vantage points:
- Standard data processing methods and their hidden capacity to solve difficult problems
- Statistical methods based on state-space models
- Methods of extracting the regularities of a process on the basis of identifying key parameters

The course addresses the problems in navigation, solar physics, geomagnetism, space weather and biomedical research and will be useful for broad range of interdisciplinary applications.

Course content and main topics

1. **Introduction to statistical analysis**
   Regression analysis and least squares methods (LSM). Limitations of LSM application in regression problems.

2. **Quasi-optimal approximation under uncertainty**
   2.1. Running mean. Classification of estimation errors and accuracy analysis
   2.2. Exponential mean. Comparison with running mean
   2.3. Applications in solar physics and biomedicine

   Short description:
   The model-independent methods are widely applied to reconstruct dynamical processes subject to uncertainties. However the simplicity of use without proper analysis of estimation error may be a trap leading to false conclusions. First part of the course covers feature analysis of processes for which simple methods provide effective solution and discuss conditions under which they break down.

3. **Optimal approximation at state space**
   3.1. Introduction to Kalman filter
   3.2. Construction of navigation filter for tracking objects
   3.3. Ill-conditioned tracking problem
   3.4. Observability and controllability
   3.5. Optimal smoothing. Forward - Backward Kalman filter
   3.6. Equivalence of exponential mean and stationary Kalman filter for the random walk model
   3.7. Optimal choice of smoothing gain
   3.8. Backward exponential smoothing and estimation accuracy increase
   3.9. Applications in navigation and solar physics

   Short description:
   The methods of filtration and smoothing based on state-space model provide optimal state estimation and estimation error. This section analyzes conditions for quasi-optimal methods to match the optimal ones. This allows increasing the utility of quasi-optimal methods under uncertainty.

4. **Process reconstruction free from any constraints and assumptions**
   4.1. Constraints and assumptions to reconstruct the process
   4.2. Broken-line and smooth curve. Optimality criterion
   4.3. Well-conditioned optimization problem
   4.4. Comparison with other approaches
   4.5. Applications in solar physics

Short description:
False assumptions about the process in question may significantly distort estimation output and lead to false conclusions. This topic covers the methods of processing the experimental data that do not need any prior assumptions about the process.

5. **Model construction at state space under uncertainty** Prior mathematical model justification
   5.1. Stochastic adaptive models. Uncertainty modeling
   5.2. Nonlinear models. Extended Kalman filter
   5.3. Noise statistics identification (state and measurement errors)
   5.4. Applications in navigation, solar physics and biomedicine

   Short description:
   One of the core problems of state estimation is the mismatch between the model and true process dynamics. Furthermore there is an uncertainty of measurement noise. One way to reduce the resulting mismatch is to use the stochastic adaptive filters on the basis of noise statistics identification.

6. **Key parameters to extract the process regularities**
   6.1. Ring analysis
   6.2. Saturation point forecasting
   6.3. Integral activity
   6.4. Others
   6.5. Applications in solar physics and geomagnetism

   Short description:
   Ability to found key parameters determining the process dynamics opens new and unexpected ways to extract hidden process regularities and improve forecasting.

Learning Outcomes:

Knowledge
1. Knowledge of state of the art data processing methods, their distinctive peculiarities and possibilities.
2. Knowledge of actual data processing problems for space applications.
3. Knowledge of approaches to estimate the accuracy and reliability of obtained results.

Skills
1. Transform theoretical knowledge into useful skills of data processing
2. Identify specific features of experimental data and choose an appropriate way of data processing
3. Detect and analyze shortcomings of data processing methods for a particular dynamical process

Experience
1. Experience in the development of estimation and forecasting algorithms of a system state in conditions of uncertainty
2. Experience in methods of process reconstruction, extraction of its tendencies and regularities
3. Experience in accuracy and reliability estimation of obtained results, prevention of false conclusions

Grading Policy for the Course
Assessment items:
1. Laboratory works
   Please note that late submissions after the deadline decrease the grade for 10% for each day of delay.
2. Exam
3. Final project
4. Attendance

All assessment items will be graded on a 100-point numerical scale and the total numerical score for the course will be calculated for each student based on the proportion of the total grade for the course allocated to each item. Each student’s numerical grade for the course will be converted to a letter according to the following schema:

Letter Grade	Meaning	Numerical Range	Explanation of meaning
A	Excellent	≥80	The work examined is outstanding and provides evidence of excellent performance demonstrating a superior understanding of the subject matter, a foundation of extensive

			knowledge, and a skillful use of concepts and/or materials. All Learning Outcomes are satisfied at a high level.
B	Good	<80 ≥70	The work is of high standard and provides evidence of comprehensive knowledge and good performance demonstrating capacity to use the appropriate concepts, a good understanding of the subject matter, and an ability to handle the problems and materials encountered in the subject. All Learning Outcomes are satisfied and a majority satisfied at a high level.
C	Satisfactory	<70 ≥60	The work examined is generally satisfactory and provides evidence of adequate performance demonstrating a sufficient understanding of the subject matter however with notable shortcomings. Majority of the learning outcomes are satisfied at an appropriate level.
D	Poor	<60 ≥50	The work examined is poor and provides evidence of very limited familiarity with the subject matter, insufficient knowledge and significant shortcomings. The evidence shows that only some of the learning outcomes were satisfied at an appropriate level.
E	Very poor	<50 ≥40	Minimally acceptable performance. The work examined is very poor, demonstrates serious deficiencies and provides little evidence of knowledge, understanding, and/or skills and familiarity with the subject matter. Very few (if any) of the learning outcomes are satisfied at an appropriate level.
F	Unacceptable (Fail)	<40	The work examined is unacceptable and provides minimum (if any) evidence of knowledge and understanding of the subject matter. The evidence fails to show that any of the Learning Outcomes are satisfied at an appropriate level.

Academic Integrity
All assignments submitted during this course must be original and prepared specifically for this course. Plagiarism is not permitted and will not be tolerated. Assignments will be checked to ensure academic integrity.

Classroom Attendance and Participation
In accordance with Skoltech's policy, **attendance in scheduled classroom sessions is compulsory.** Regular attendance, and active participation in classroom discussions, will be an important part of the learning process, and will help you to contribute appropriately to the work of your team. Special permission from the instructor will be required to be excused from a classroom session. Penalties will apply to students who miss class without permission. **Please note that penalties from missing class without permission will be serious enough to have a substantive effect on your final letter-grade for the course.**

Schedule

Date	Day of the week	Topic
October 1	Monday	Introduction to data processing Topic 1. Introduction to statistical analysis

October 2	Tuesday	Laboratory work 1. Relationship between solar radio flux F10.7 and sunspot number
October 3	Wednesday	No class Industry day
October 4	Thursday	Topic 2. Quasi-optimal approximation under uncertainty Laboratory work 2. Comparison of exponential and running mean for random walk model I. Determination of optimal smoothing constant in exponential mean. II. Comparison of methodical errors of exponential and running mean.
October 5	Friday	Discussion of laboratory work 3 Laboratory work 3. Determining and removing drawbacks of exponential and running mean. Task 1. I. Backward exponential smoothing. II. Drawbacks of running mean
October 8	Monday	Laboratory work 4. Determining and removing drawbacks of exponential and running mean. Task 2. I. Comparison of the traditional 13-month running mean with the forward-backward exponential smoothing for approximation of 11-year sunspot cycle. II. 3d surface filtration using forward-backward smoothing.
October 9	Tuesday	Topic 3. Optimal approximation at state space Laboratory work 5. Tracking of a moving object which trajectory is disturbed by random acceleration
October 10	Wednesday	Laboratory work 6 Analysis of accuracy decrease of tracking in conditions of biased state and measurement noise I. Divergence of tracking filter when bias of state noise is neglected in assimilation algorithm. Development of optimal Kalman filter that takes into account bias of state noise
October 11	Thursday	Laboratory work 7. Development of optimal smoothing to increase the estimation accuracy
October 12	Friday	Development of tracking filter of a moving object when measurements and motion model are in different coordinate systems Laboratory work 8. Tracking and forecasting in conditions of measurement gaps

October 15	Monday	Laboratory work 9. Development of tracking filter of a moving object when measurements and motion model are in different coordinate systems I. Instability zone of a tracking filter due to ill-conditioned coordinate transformations of measurements.
October 16	Tuesday	No class MIT-Skoltech Conference
October 17	Wednesday	Topic 4. Process reconstruction free from any constraints and assumptions Laboratory work 10. Process reconstruction free from any constraints and assumptions
October 18	Thursday	Topic 5. Model construction at state space under uncertainty I. Extended Kalman filter for navigation and tracking Laboratory work 11. Extended Kalman filter for navigation and tracking
October 19	Friday	Laboratory work 12. Joint assimilation of navigation data coming from different sources
October 22	Monday	Laboratory work 13. Vehicle tracking based on GPS and odometry data fusion
October 23	Tuesday	Topic 6. Key parameters to extract the process regularities. Summary and final project discussion.
October 24	Wednesday	Exam
October 25	Thursday	Evaluation period, work on final project/presentations
October 26	Friday	Application period, project submission