Assignment 4
Optimization & Real Data Analysis

Deadline:	Tuesday December 9th, 2014, 23:59h	
Subjects:	Optimization techniques	Lect. 09
	Physical modelling	Lect. 10
	Parameter accuracy	Lect. 11

Important notes
- Reports handed in after the deadline will not be marked, resulting in a zero.
- Working in teams of (not more than) two students is strongly encouraged.

How to get help
- Always check the *Assignment Guide* on Blackboard for support.
- The command functions *HELP* or *DOC* <*FUNCTIONNAME*> can be used for instant support to MATLAB functions.
- Assistance hours (see schedule).
- Help each other using the *Discussion Board* on Blackboard.

Report requirements
- Reports should be written in English.
- Each student hands in a report using blackboard. Always provide the name of the partner at the report.
- Always provide explanations in your answers and be precise and concise (providing irrelevant information can negatively influence your mark). Formulate your answer in such a way that you show you understood the question.
- Use figures to illustrate your findings. If possible use the same figure for the results of multiple sub-questions; this allows for easy comparison. When plotting more than one curve in a figure, use legends. Do not forget a caption and the units on the axes. Choose proper axes scales. Unclear Figures and Tables will be excluded from judgment.
- Do not submit MATLAB coding, unless it is clearly stated that you should.
- Put your name (and that of your colleague if applicable), page number, and the assignment number on each page of your report.

Goal
Learn how to use different optimization techniques to find a reduced set of parameters.

Accessories:

Question 1:	ASS4.M	**Question 2:**	DATAREC1.MAT
	DATA1.MAT		DATAREC2.MAT
	ERRFUN.M		ERRFUNMBK.M
	CUSTOMPLOTFCN.M		VAF.M
	STOREPATHWAY.M		MODMBKKV.MDL
			FREQAVG.M

Question 1: Optimization Techniques

Open *ASS4.M* and take a look at *DATA1.MAT*. The data contains a time vector (*t*) and a measured data vector (*y*). Using prior knowledge of the measured system, we define a model-based relationship between the input (*a* & *b*) and *y* as:

$$\hat{y}(t) = a\cos(\tfrac{1}{2}b \cdot t) + b\sin(\tfrac{1}{5}a \cdot t)$$

With *a* & *b* within the range of [0:10]. The goal is to find the correct input parameters for the model by optimization (minimization of the error between the measured data and the model output), which describes the measured data as good as possible. Apply the following optimizations:

- <u>Grid Search:</u>
  Use a Grid Search (resolution 0.1) for optimization of the model. Define *ERRFUN.M* with as output the model output (*ymod*) and the summed squared error between *y* and *ymod* are defined. Plot the summed squared error as function of *a* and *b (use function surfc)*.
- <u>Gradient Search:</u>
  Use a Gradient Search for optimization of the model. Build your own *ERRFUN2.M* with as output the model output (*ymod*) and vectorized (unsquared) error between *y* and *ymod* are defined (note that the used Matlab optimization takes the sum-of-squares of the vector!). Use a=b=5 as initial conditions. Also try several other initial conditions. Plot the pathways on the surface plot of the grid search.
- <u>Genetic Algorithm:</u>
  Use a Generic Algorithm for optimization of the model.

a) Include the plots of the grid search and the gradient search in your report and comment on these graphs.

Make a table of the results of all optimization techniques containing resulting input values (i.e. the optimal a & b), error residual (summed squared), number of iterations, time spend on calculation (*use functions tic and toc*) and discuss these results. (Use for the grid search the mentioned resolution and for the gradient search the suggested initial conditions).

b) Include the table in your report and give at least one advantage and one disadvantage for all techniques. Keep in mind that normally much more complex models are used (with more variables/inputs).
c) How could we improve the estimate for the grid search? Motivate your answer. What are the (dis)advantages of your improvement?
d) Look into the effect of all options of the genetic algorithm (*POPULATIONSIZE*, *ELITECOUNT*, *CROSSOVERFRACTION*, *POPINITRANGE* & *GENERATIONS*) by performing several optimizations. Describe shortly their purpose and what happens if you choose them wrong.
e) Suggest and implement a combination of grid search and gradient search which **robustly** finds the global minimum independent of initial position in less function evaluations than a full grid search. Motivate your answer and include the Matlab code.

Question 2: Real Data Analysis

In this question you will analyse a real data set recorded from the human wrist. In the experiment the wrist is perturbed with small angle rotations around a neutral position (0 rad) while the subject is requested to maintain constant levels of torque in a stair-like pattern (look at the movie). Measurements are made of torque and angle. We want to use this data to see how mechanical parameters describing the human wrist joint change with respect to exerted torque. The bandwidth of the applied perturbations is 2-20 Hz to discern between 'slow' voluntary changes in torque and reaction torques driven by the perturbations. Frequency spacing is chosen such that segments of data can be analysed separately without leakage.

__Frequency Domain:__

Run part A of *ASS4.M*

a) Figure(1) displays the raw data. Add the appropriate axes.
b) Figure(2) shows the frequency response function for the whole dataset and for separate data segments with a constant torque level. Give two reasons why the coherence for the whole dataset is smaller than for the separate data-segments. Support your reasoning with a plot (or two).
c) Is it wise to apply linear time invariant (LTI) identification to this dataset as a whole? Explain.
d) Explain how the dynamic characteristics can be used to estimate parameters of a model in the frequency domain. In what frequency ranges do the mass, the spring and the damper affect the frequency response function (the most)?
e) Part B can be used to estimate the mass, spring and damper parameter by using lsqnonlin.m. Do the estimated transfer functions fit the dynamics well? Explain why.
f) Improve your estimates by only estimating on the frequencies which are perturbed and by implementing frequency and coherence weighing in the error-function errfunMBK.m.

Note: LSQNONLIN comes with a lot of options which do heavily affect performance. Check the significance of at least the following options and change them if necessary: 'TolX', 'TolFun', 'DiffMinChange', 'MaxFunEval'.

If you succeeded with f) you have a reasonable model now, although clearly something is still missing. We know that humans use afferent feedback for motion control. Supplement your model with velocity feedback. The feedback should contain a feedback gain '*Kv*', a time delay '*td*' and a second order transfer function representing activation dynamics. Implement this activation filter with one free parameter '*w*' representing the cut-off frequency:

$$H_{act} = \frac{w^2}{s^2 + 2 \cdot 0.7 \cdot w \cdot s + w^2}$$

g) Derive the complete transfer function with velocity feedback in the s-domain and include it in your report.

h) Create errfunMBKKv.m with your new transfer function and with a set of 6 parameters to be optimized: [M B K Kv td w]. Now use lsqnonlin to estimate these parameters. Initial values and lower/upper boundaries are given in part C of Ass4.m. Give a plot of your new estimates

<u>Validation & Time Domain:</u>

So far, you have been estimating parameters without having the means of quantitatively evaluating the accuracy of the parameters and the goodness of the fit. Now we will be using two such measures: the standard error of the mean (SEM) and the variance accounted for (VAF).

i) Look at the SEM-values of your latest estimate. Which of the parameters are estimated less accurate? Give a plot to support your result.
j) Which parameters change with background torque and which do not? Can you relate this to the underlying physiology?
k) Use Part D to create the time-domain response and give a table of the VAF values for each segment. Compare the estimated and measured rotation by eye and relate this to the VAF value.

Now we are going to try to fit the parameters in the time-domain. Something to start with is already made for you in part E. Estimates are improved if we first get rid of the low-frequent drift in the torque signal using a high-pass

l) What would happen with your estimates if you didn't use the high-pass filter?
m) Write an error function (start off with errfunMBKKv.m) using the simulation as done in part D, and estimate the parameters in the time-domain.
n) Compare results of the frequency and time domain estimates in terms of VAF, SEM and parameter values. Discuss differences and similarities. *Hint: For proper comparison redo k) with high-pass filtered data*
o) In the time-domain we did not get rid of the high frequent noise (as we did in the frequency domain by limiting the frequency band for estimation) but still got a good fit. Explain why.
p) Now use the second dataset to determine the VAF using parameters estimated on the first dataset. Why is it better to use another dataset for validation?
q) Give an advantage and disadvantage of time-domain identification with respect to frequency-domain identification.