Abstract

This document contains relevant information about the IML exam. The document is quite accurate but tentiative until mid december.

Exam Preparation

Dan Witzner Hansen

December 2025

Figure 1: Map of the IML content.

Exam

The exam is individual, oral and will last up to 30 minutes.

1 Before Exam

We urge you to thoroughly prepare for the exam. This includes:

1. Read all material, tutorials and watch the video material (the curriculum).
2. Prepare and practice a presentation for each exam question and its corresponding focus point (listed as bullet points below)
   (a) **Prioritization of Exercises:** Ensure you have covered the *"red"* (expected) exercises from the exercises. These are essential and

should form the core of your presentation including some theory, methods and results. You may weave medium- and low-priority topics into your presentation to demonstrate you understanding.

(b) **Medium and Low Priority Exercises:** While completion of *"yellow"* (medium priority) and *"green"* (low priority) exercises is not expected, you are encouraged to integrate aspects of these topics into your presentation where appropriate to provide additional context or depth. Focus on elaborating the relevant theory when encountering questions that entail medium and low prioroty questions. Implementation details and exercise results are not required.

3. You are **required** to bring your Jupyter notebooks and examples to the exam. These should be used to illustrate outcomes and discuss their significance in relation to the exam question. Additionally, you are permitted to bring a summary outline (which you can conveniently incorporate into the Jupyter notebook).

   While you may display the results of your exercises, **avoid** demonstrating or executing code during the exam unless explicitly requested. This will allow you to focus on discussing the issue at hand, including its theoretical context, formal models, data, and findings.

2 At the Exam

A video has been made available on LearnIT that illustrates the following steps.

Enter the room Be ready in in front of the the examination room (or in zoom waiting room) at least 60 minutes before your scheduled time.

Draw Question (0-1') You will randomly select (rolling dice) a question corresponding to one of the weekly exercises or mandatory assignments. Each question includes multiple focus points, and one will be chosen randomly by rolling a set of dice.

Present Question (1-6') In 5-6 minutes, you **must** present the exercise, including an overview, purpose, main findings, and a bit of theoretical context. Given the limited time, focus on the most substantial aspects of the exercise, using the provided focus points to demonstrate depth and relevance to the course material.

Answer General Questions (6-25') Initially, we will not interrupt your presentation, but we may ask clarifying questions. As time progresses, we will pose questions about your exam question, theory, and other topics within the curriculum. Be aware that the course curriculum includes material not encompassed by the exercises.

Evaluation (25-30') After about 20-25 minutes, your direct examination ends, and you will be asked to wait outside the room (in the waiting room if online). You must take your computer and other belongings with you outside the examroom. After deciding on a grade we will invite you back to inform you of your grade with brief feedback. Detailed feedback may be scheduled for another day upon request.

Criteria for Assessment You are not assessed on your ability to recite your exercise implementation (do not describe the code details), but on your capacity to describe and reason about the methods accurately, connect it to the underlying theory, and demonstrate analysis, interpretation, and critical reflection; simply repeating formulas, definitions, plots, or procedural steps is not sufficient.

To achieve an average (4-7) or higher grade, you must provide, in addition to an accurate description of the exercise (or how it could be implemented), a deeper discussion of the machine learning models applied, the rationale for their use, and their relation to the underlying theory and mathematics. An excellent grade (10-12) requires precision together with deeper theoretical insight and critical evaluation and goes beyond description by demonstrating analysis, interpretation, and critical reflection.

For the descriptive parts of the student's answer, emphasis will be placed on how clearly and correctly the exercise is presented, including the steps taken and the results obtained. While the exact recitation of formulas is not in itself central, you should be able to explain the purpose of key formulas, how they operate, and how they are used in practice. Formal proofs are not required, but it is important that students can provide clear intuition and describe the mathematical role of the methods and concepts they use. Furthermore, we will assess the student's ability to independently select and prioritize the most relevant aspects of the syllabus in relation to the exercise.

Questions

The bullet points listed for each week denote the specific focus areas for the corresponding questions. At the exam, one of these bullet points will be selected at random (rolling dice) to determine your precise focus for discussion (see video for details)

Q1 Exercises Week 2-3 (Vector & Matrices)

(a) Focus on inner products, vector operations, distance metrics and their relation to ML (evaluation and other metrics). You may relate this to week 10 (evaluation) but focus on vectors.

(b) Focus on matrices and their operations (addition, multiplication, transpose, inverse, determinant, orthogonal). Relate them to their application within the course (e.g., transformations, basis in week 11).

(c) Focus on how linear equations and their solutions are related to matrices.

Q2 Exercise Week 4 (Linear Transformations)

(a) Use the tutorial to focus on exploring linear transformations in 2D and 3D spaces, including operations such as scaling, shearing, reflections, rotations, and translations, while drawing connections to their extensions in higher-dimensional linear transformations. Additionally, you should explain the relationship between linear transformations and non-linear transformations including affine, thus bridging the gap between linear and more complex transformations.

(b) Focus on how linear and non-linear models can be learned using matrix inverses. Discuss the relationship between model complexity (e.g., polynomial degree) and the amount of data needed to accurately train these models with matrix inverses.

(c) Focus on affine transformations, homogeneous coordinates and composition of linear transformations.

Q3 Exercise Week 5 (Projections and Least Squares)

(a) Focus on the relation between linear least squares (function minimization) and projections.

(b) Focus on linear least squares problems for model fitting (design matrix, kernel, lines, polynomials, affine, and other multivariate functions) and the interpretation of results for various types of models (see week 7).

(c) Learning of Affine (multivariate) functions and linear optimization.

Q4 Exercise Week 6 (Mandatory 1)

(a) Focus on preprocessing and feature extraction in Mandatory 1

(b) Focus on model predictions and learning Mandatory 1

(c) Focus on model evaluation in Mandatory 1

(d) Describe vector space, basis, independence and how these are related to concepts in machine learning (learning, transformations etc). You MAY include exercises week 12 if you have done them.

Q5 Exercise Week 7 (Model Complexity, Model seleection and noise)

(a) Linear least squares problems for model fitting (design matrix, kernel, lines, polynomials, affine, and other multivariate functions).

(b) Discuss the relationship between data quality, uncertainty, and the challenges of overfitting and underfitting in model learning.

(c) Discuss how factors affect a model's ability to generalize to unseen data. Relate this to concepts of data evaluation, such as train-test splits, cross-validation, and performance metrics, to assess a model's predictive capability and robustness.

Q6 Exercise Week 8 (Filtering)

(a) Focus on filtering (1D, 2D, and n-D convolution and correlation, blurring and smoothing, and noise).

(b) Focus on filtering for derivatives, gradients, and edges. You may include how these operations can be used to construct features such as HOG.

Q7 Exercise Week 9 (Understanding Data and Descriptive Methods)

(a) Describe correlation and covariance, explain how they are calculated, and discuss how they are useful for understanding data and for descriptive analysis.

(b) Explain how descriptive statistics can be used to assess model selection and the generalization ability of machine learning models.

(c) Describe noise, outliers, and missing data.

(d) Focus on uncertainty, noise, data cleaning in relation to regression, classification, clustering or dimensionality reduction (Covariance, distributions (e.g Normal/Gaussian))

Q8 Exercise Week 10-11 (Classification)

(a) Linear classification, kernels, and classification boundaries

(b) Logistic regression and classification boundaries

(c) Linear and non-linear decision boundaries including SVM (include week 10) and possibly HOG features.

Q9 Exercise Week 11 (Evaluation)

(a) Metrics/Evaluation of Classifiers
(b) Metrics/Evaluation of Regression
(c) Imbalanced data for classification and regression

Q10 Exercise Week 12 + Assignment 2 (Principal Component Analysis)

(a) Basis and transformations
(b) Dimensionality reduction and PCA. Focus on mandatory 2
(c) Generating models and PCA
(d) Eigenvalues, covariance matrix and basis

Q11 Exercise Week 13 (Clustering)

(a) K-means and Mean shift
(b) K-means and Algomerative clustering
(c) Kmeans and ELBOW

Q12 Exercise Week 14 (Neural networks)

(a) Neural networks prediction (regression vs classification)
(b) Neural networks training (Gradients, the chain rule and back/forward propagation)
(c) Training and Evaluation including over/underfitting
(d) Model architectures: Difference between fully connected /multi layer perceptron (MLP) and CNN
(e) Loss functions, model complexity, cross-validation.

3 Change of Schedule

Early changes to schedule. An early exam plan will be available before Christmas, with only minor changes expected thereafter.While we aim to follow the schedule, some adjustments may still occur. **Inform Dan as early as possible if any circumstances, such as illness, prevent you from attending the exam.** This will ensure a fair and timely schedule for all. Once the final schedule is set (before Christmas), changes can only be made by arranging a swap with another student. To do so, coordinate with the other student, agree on the swap, and email Dan (cc'ing all parties) to notify him of the change. Please note that currently unavailable slots cannot be used for rescheduling. While uncommon, individual exam dates may still be adjusted up to one week before the scheduled exam day.

Changes on the exam date. It is essential that you check the most up-to-date exam schedule on the morning before your exam to confirm any recent changes. If a student scheduled as the next in line is absent on the day of the exam, we will proceed by calling the following student. Please be aware that there may be delays, and as a result, the scheduled exam times might shift throughout the day. Therefore, we ask that you be available and ready during the morning or afternoon session (at least 60 minutes before), as per your assigned schedule.

Additional preparation

The subsequent subsections present reflective questions designed to engage you with key elements of the course. This document is evolving and may be updated. While not exhaustive, these questions offer a means to contemplate various aspects of the course content. They are phrased broadly to facilitate a wide range of reflections.

4 Linear Algebra

Basics

- List as many purposes for which we use vectors for image analysis and Machine learning
- What is the equation of a line, planes and hyperplane using vector notation?
- How do you calculate the length and orientation of a vector?
- How do we know when two vectors u and v are orthogonal to each other?
- How do we know when two vectors u and v are parallel to each other?

Linear Equation

What is a linear equation and how is this relate to matrices?

Inner product

How is the inner product related to:

- a measure of distance.
- matrix multiplication
- projections
- convolution
- neural networks

Solutions to Linear Equation

- What does it mean to have a solution to a linear set of equations?
- When can we have one, zero or many solutions to a linear set of equations?
- What is an over-determined set of equations.
- What is an under-determined set of equations.
- Why is the Determinant relevant when talking about solutions to linear equations.
- Why are subspaces important when talking about solutions to linear set of equations.
- Given data $X \in \mathbb{R}^N$ and labels y. How w can linear equations be used to find the coefficients of the following models and how much training data is needed to learn the model parameters
  (a) A straight line in the plane
  (b) A plane in 3D
  (c) A hyperplane in N-dimensional spaces
  (d) Find the coefficients of a an N-order polynomial
  (e) Find the coefficients of an similarity or affine transformations
- In the above cases what is the minimal number of points needed to solve the linear set of equations.

Transformation

- What is a transformation and how is it related to a projection.
- Matrix multiplications may be considered as a transformation. Why?
- How are linear transformations combined?
- What is the purpose of homogeneous coordinates.
- What is the inverse of a transformation and what is its relevance to the course / ML.
- How is least squares (formally) related to projections

5 Signals

- What are the definitions of convolution and correlation and how are they related
- When can correlation and convolution be used interchangeably
- How can correlation be implemented in a neural architechture and why is this beneficial?

- How is image templates useful as a machine learning model and a metric for comparison.
- How do image templates relate to machine learning

6 Machine Learning

- Where is supervised and unsupervised learning used and how do they differ.
- How is least squares used in machine learning and how does it relate to least squares when using matrices.
- What is an objective function
- Which methods can be used to learn linear and non-linear models
- what is the difference between linear and non-linear models (including affine)
- What is the difference in how to optimize/minimize linear and non-linear functions (such as a loss function)
- Describe cases in which a non-linear prediction model can be learned with linear optimization methods
- How do recommender systems work and how is this related to inner products and matrix factorization.
- What is PCA? and how does PCA make use of subspaces , eigenvalues and eigenvectors.
- How is PCA and certain neural architectures related.
- Why is it called linear classification
- How does logistic regression differ from linear classifiation.
- What is the decision boundary and how can you find it.
- what is a kernel, where are they used and how is it related to model learning

7 Evaluation

- Why is evaluation needed
- How do you ensure proper evaluation of models
- Why are training, test and verification sets needed in the training procedures?

- What is cross validation and how is it related to overfitting / underfitting
- How how can we tell when a model is under and overfitted?