LABORATORY MANUAL

ECE 403/503

OPTIMIZATION for MACHINE LEARNING

This manual was prepared by

Wu-Sheng Lu

University of Victoria
Department of Electrical and Computer Engineering

A. SAFETY REGULATIONS

Safe practice in the laboratory requires an open attitude and knowledgeable awareness of potential hazards. Safety is a collective responsibility and requires full cooperation from everyone in the laboratory. This cooperation means each student and instructor must observe standard safety precautions and procedures and should:

1. Follow all safety instructions carefully.
2. Become thoroughly acquainted with the location and use of safety facilities such as fire extinguishers, first aid kits, emergency showers, eye- wash stations and exits. See marked floor plan posted near exit staircases for equipment locations.
3. Become familiar with experimental procedures and all potential hazards involved before beginning an experiment.
5. Be sure line-powered equipment has 3-wire grounding cords and that you know how to use the equipment properly. Ask for help and instruction when needed.
6. Any electrical failure or evidence of undue heating of equipment should be reported immediately to the instructor and/or technologist. If you smell over-heating components or see smoke coming from any circuit or equipment, switch the power off immediately.
7. Ensure all equipment is powered-off at the end of each experiment.
8. Only qualified ECE personnel should maintain electric or electronic equipment.
9. Cardiopulmonary resuscitation (CPR) often will revive victims of high-voltage shock. Only qualified people should attempt CPR.

ELECTRICAL HAZARDS

General Safety Principles

Electrical currents of astonishingly low amperage and voltage under certain circumstances may result in fatal shock. Low-voltage dc circuits do not normally present a hazard to human life, although severe burns are possible. Voltage as low as 24 volts ac can be dangerous and present a lethal threat. The time of contact with a live circuit affects the degree of damage, especially where burns are concerned. Very small electrical shocks, even small "tingles", should serve as a warning that an electrical problem exists and that a potentially dangerous shock can occur. The equipment and circuits in use must be immediately disconnected and not re-used until the problem is discovered and corrected by the instructor and/or technologist.

1. When handling electric wires, never use them as supports and never pull on live wires.
2. Report and do not use equipment with frayed wires or cracked insulation and equipment with damaged plugs or missing ground prongs.
3. Report and do not use receptacles with loose mountings and/or weak gripping force.
4. Avoid pulling plugs out of receptacles by the cord and avoid rolling equipment over power cords.

FIRST AID

1. For simple cuts or minor first aid use the First Aid Kits available in each room. The University Health Services may also be contacted at **8492**. All injuries, no matter how minor, should be reported to your instructor and/or technologist.
2. For medical emergencies call **911** <u>and</u> Traffic and Security at **7599**.

INDIVIDUAL RESPONSE PROCEDURES FOR FIRES

If **<u>you discover</u>** a fire, smoke or an explosion:

1. Shout for assistance.
2. Activate the nearest fire alarm.
3. If it is a small fire, attempt to put it out with available fire equipment. See the marked floor plan posted near exit staircases for equipment locations.
4. If the fire is out of control and it is too large to handle with one fire extinguisher, isolate the fire by closing the doors and windows behind you as you leave. Do **not** lock the doors.
5. Warn others and leave the building with reasonable speed using recommended exits. Assist disabled and injured persons in reaching assembly areas when conditions permit.
6. Stand by to identify yourself and provide information to fire personnel.

If a **Fire Alarm sounds**:

1. Secure any equipment you are using and switch the power off.
2. Close windows and doors behind you as you leave. Do **not** lock doors.
3. Leave the building with reasonable speed using recommended exit.
4. Follow instructions of your floor warden or deputy. Wardens will ensure evacuation of assigned rooms.
5. **DO NOT** use elevators for evacuation.
6. **DO NOT** re-enter the building until allowed to do so by the Fire Department.

INDIVIDUAL RESPONSE PROCEDURES FOR EARTHQUAKES

IF INDOORS:

Take action at the first indication of ground shaking.

1. Stay inside; move away from windows, shelves, heavy objects and furniture that may fall. Take cover under a table or a desk, or in a strong doorway (anticipate that doors may slam shut).
2. In halls, stairways or other areas where no cover is available, move to an interior wall. Turn away from windows, kneel alongside the wall, bend your head close to your knees, clasp your hands firmly behind your neck covering the sides of your head with your elbows.
3. Elevators must not be used. They are extremely vulnerable to damage from earthquakes. Ground shaking may cause counterweights and other components to be torn from their connections, causing extensive damage to the elevator cabs and operating mechanisms.
4. When exiting a building, move quickly through exits and away from buildings. Parapets and columns supporting roof overhangs may fall.
5. Assemble away from gas, sewer and power lines.

IF OUTDOORS:

1. Move to an open space away from buildings, trees and overhead power lines.
2. Lie down or crouch low to the ground (legs will not be steady) and constantly survey the area for additional hazards.

B. LABORATORY OPERATION GUIDELINES

During the operation of the laboratories, the following simple procedures and guidelines are essential and **must** be adhered to by all students.

- **FOOD, DRINKS and SMOKING** are **NOT** permitted in the laboratories.
- Before starting your experiment, make sure proper equipment and circuit connections are made as per instructions in the laboratory manual. Verify your set-up with your instructor.
- All damaged or missing equipment or parts should be reported as soon as possible to the technologist. A Repair Request form (available in each lab) must be completed before equipment can be serviced.
- Equipment should not be removed from the lab station. If equipment is required elsewhere, it is to be returned to the lab station once the requirement is finished.
- All electronic components, such as capacitors, resistors, transistors etc., must be returned to their respective storage trays when the experiment is finished.
- All leads and cables are to be returned to the wall racks when the lab is finished and oscilloscope probes are to remain with the oscilloscope.
- Benches and equipment set-ups are to be tidied up after each lab session. All garbage is to be placed in the garbage cans provided. No writing on equipment or benches is permitted.
- If students have any questions about the experiment, they should consult the instructor first and then the technologist.
- Abide by all safety rules and regulations of the laboratory.

ELECTRICAL AND COMPUTER ENGINEERING LABORATORIES

ACKNOWLEDGMENT FORM

NAME: ______________________________

COURSE: ______________________________

LAB INSTRUCTOR: ______________________________

Experiments conducted in the Electrical and Computer Engineering laboratories involve the handling of electric and electronic equipment and circuits. Failure to handle this equipment properly may lead to injury or even fatal shock. For the safety of everyone, it is required that you understand and follow the appropriate laboratory procedures as outlined in the laboratory manuals and by your laboratory instructor.

Your signature below is your acknowledgment that you have read the *Safety Regulations* and the *Laboratory Operation Guidelines* and agree to abide by them.

______________________________ ______________________

STUDENT’S SIGNATURE DATE

____

STUDENT NUMBER

Preparing Your Laboratory Report

The objective of the experiments described in this manual is to familiarize the student with computer simulations and implementation of several optimization as well as data processing techniques as they are applied to machine learning problems. The primary software tool required in all experiments is MATLAB to which the student can access during the laboratory sessions. An appendix, that introduces main functions in MATLAB and their usage, is included to facilitate the students in preparing their MATLAB code.

PREPARATION

Successful completion of an experiment depends critically on error-free MATLAB programming. Therefore, preparation prior to the laboratory period is essential. Specifically, the student should *study the description of the experiment and prepare useable MATLAB codes required in the preparation section(s)* ***before*** *the experiment is carried out*. The student will be required to present the preparation at the beginning of the lab session.

THE LABORATORY REPORT

A laboratory report is required from each group for each experiment performed. The lab report should be submitted **within one week after the experiment**. The front page of the report is shown on the next page and should be used for each laboratory report.

The report should be divided into the following parts:
(a) Objectives.
(b) Introduction.
(c) Results including relevant MATLAB programs and figures, and description of the implementations.
(d) Discussion.
(e) Conclusions.

UNIVERSITY OF VICTORIA
Department of Electrical and Computer Engineering

ECE 403/503 Optimization for Machine Learning

LABORATORY REPORT

Experiment No:

Title:

Date of Experiment:
(should be as scheduled)

Report Submitted on:

To:

Laboratory Group No.:

Name(s): (please print)

Experiment 1

Handwritten Digits Recognition Using PCA

1. Introduction and Objectives

Research for handwritten digit recognition (HWDR) by machine learning (ML) techniques has stayed active in the past several decades. The sustained interest in HWDR is primarily due to its broad applications in bank check processing, postal mail sorting, automatic address reading, and mail routing, etc. In these applications, both accuracy and speed of digit recognition are critical indicators of system performance. In a machine learning setting, we are given a training data set consisting of a number of samples of handwritten digits, each belongs to one of the ten classes $\{\mathcal{D}_j, j = 0, 1, \ldots, 9\}$ where class $\mathcal{D}_j$ collects all data samples labeled as digit j. The HWDR problem seeks to develop an approach to utilizing these known data classes to train a multi-category classifier so as to recognize handwritten digits *outside* the training data. The primary challenge arising from the HWDR problem lies in the fact that handwritten digits (within the same digit class) vary widely in terms of shape, line width, and style, even when they are properly centralized and normalized in size.

Classification of multi-category data can be accomplished by ML methods that are originally intended for classifying binary-class data using the so-called *one-versus-the-rest* approach. Alternatively, there are ML techniques that deal with multi-category data directly. A representative method of this type is based on principal component analysis (PCA). In Sec. 1.3 of the course notes, PCA is introduced as a means for feature extraction. The objective of this laboratory experiment is to learn PCA as a technique for pattern recognition and apply it to the HWDR problem.

2. The Dataset

The database MNIST for HWDR was introduced in Example 1.1 of Section 1.2 of the course notes. A full-scale training data from MNIST fit into the structure $\mathcal{D} = \{(\boldsymbol{x}_n, y_n), n = 1, 2, ..., N\}$ with $N = 60{,}000$, where, for each digit $\boldsymbol{x}_n$, a label is chosen from $y_n \in \{0, 1, ..., 9\}$ to match what $\boldsymbol{x}_n$ represents. Originally each $\boldsymbol{x}_n$ is a gray-scale digital image of 28×28 pixels, with components in the range $[0, 1]$ with 0 and 1 denoting most white and most black pixels, respectively. In this experiment, each $\boldsymbol{x}_n$ has been converted into a column vector of dimension $d = 784$ by stacking matrix $\boldsymbol{x}_n$ column by column. In this way, each digit from MNIST can be regarded as a "point" in the 784-dimensional Euclidean space. If we put the entire training data together, column by column, to form a matrix $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_m \end{bmatrix}$, then $m =$ 60,000 and $\boldsymbol{X}$ has a size of 784×60000.

3. PCA for Multi-Category Classification

In the HWDR problem, we are given a training set that consists of 10 data classes of the form $\mathcal{D}_j = \{(\boldsymbol{x}_i^{(j)}, y_j), i = 1, 2, ..., n_j\}$ for $j = 0, 1, \ldots, 9$, where all samples $\{\boldsymbol{x}_i^{(j)}\}$ in class $\mathcal{D}_j$ are associated with the same label $y_j = j$. Our goal is to classify (i.e. recognize) a new digit $\boldsymbol{x}$ *outside* the training data as one of 10 possible digits. To do so, PCA is applied to *each* data class $\{(\boldsymbol{x}_i^{(j)}, y_j), i = 1,2,\ldots,n_j\}$ by first computing the mean vector and covariance matrix as

$$\boldsymbol{\mu}_j = \frac{1}{n_j}\sum_{i=1}^{n_j}\boldsymbol{x}_i^{(j)}, \quad \boldsymbol{C}_j = \frac{1}{n_j}\sum_{i=1}^{n_j}(\boldsymbol{x}_i^{(j)} - \boldsymbol{\mu}_j)(\boldsymbol{x}_i^{(j)} - \boldsymbol{\mu}_j)^T \quad \text{for } j = 0, 1, \ldots, 9 \tag{E1.1}$$

and then performing eigen-decomposition of $\boldsymbol{C}_j$ as

$$\boldsymbol{C}_j = \boldsymbol{U}_j\boldsymbol{S}_j\boldsymbol{U}_j^T \quad \text{for } j = 0, 1, \ldots, 9 \tag{E1.2}$$

Next, we construct a rank-q approximation for each $\boldsymbol{C}_j$ as

$$\boldsymbol{C}_j \approx \boldsymbol{U}_q^{(j)}\boldsymbol{S}_q^{(j)}\boldsymbol{U}_q^{(j)T}$$

where

$$\boldsymbol{U}_q^{(j)} = [\boldsymbol{u}_1^{(j)}\ \boldsymbol{u}_2^{(j)} \cdots \boldsymbol{u}_q^{(j)}] \quad \text{for } j = 0, 1, \ldots, 9 \tag{E1.3}$$

consists of q eigenvectors of $\boldsymbol{C}_j$ corresponding to the q largest eigenvalues and $\boldsymbol{S}_q^{(j)}$ is a diagonal matrix that collects the q largest eigenvalues in descent order. So now we get 10 pairs $\{\boldsymbol{\mu}_j, \boldsymbol{U}_q^{(j)}\}$ for $j = 0, 1, \ldots, 9$, where $\boldsymbol{\mu}_j$ denotes the average character of the jth class $\mathcal{D}_j$, and $\boldsymbol{U}_q^{(j)}$ are q principal axes that define a q-dimensional subspace in which the variations of the individual members in the jth class from its mean $\boldsymbol{\mu}_j$ can be well represented. The dimension reduction becomes substantial when $q \ll d$. It is important to note that the quantities described above can be prepared off-line prior to the real-time task of classifying a new sample $\boldsymbol{x}$. The classification of a point $\boldsymbol{x}$ is carried out in the following steps:

(i) Project point $\boldsymbol{x} - \boldsymbol{\mu}_j$ into the jth data class for $j = 0, 1, \ldots, 9$. These projections are performed by computing principal components

$$\boldsymbol{f}_j = \boldsymbol{U}_q^{(j)T}(\boldsymbol{x} - \boldsymbol{\mu}_j) \quad \text{for } j = 0, 1, \ldots, 9 \tag{E1.4}$$

(ii) Represent point $\boldsymbol{x}$ in the jth class as

$$\hat{\boldsymbol{x}}_j = \boldsymbol{U}_q^{(j)}\boldsymbol{f}_j + \boldsymbol{\mu}_j \quad \text{for } j = 0, 1, \ldots, 9 \tag{E1.5}$$

(iii) Compute the Euclidean distance between $\boldsymbol{x}$ and its representation $\hat{\boldsymbol{x}}_j$ as

$$e_j = \|\boldsymbol{x} - \hat{\boldsymbol{x}}_j\|_2 \quad \text{for } j = 0, 1, \ldots, 9 \tag{E1.6}$$

(iv) Sample $\boldsymbol{x}$ is classified to class j^* if e_{j^*} reaches the minimum among $\{e_j, j = 0, 1, \ldots, 9\}$, that is,

$$\text{sample } \boldsymbol{x} \text{ is classified to class } j^* = \arg\min_{j=0,1,\ldots,9}\{e_j\} \tag{E1.7}$$

Below we summarize the complete procedure as an algorithm.

PCA-Based Multi-Category Classification Algorithm

Input: Training data that contains 10 classes $\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_9$; the number q of principal axes to be used; and a new data point $\boldsymbol{x}$ to be classified.

Step 1 Compute $\boldsymbol{\mu}_j$ and $\boldsymbol{C}_j$ using (E1.1) for each data class $\mathcal{D}_j$, for $j = 0, 1, \ldots, 9$.

Step 2 Compute the q eigenvectors of $\boldsymbol{C}_j$ corresponding to the q largest eigenvalues and use them to construct matrix $\boldsymbol{U}_q^{(j)} = [\boldsymbol{u}_1^{(j)}\ \boldsymbol{u}_2^{(j)} \cdots \boldsymbol{u}_q^{(j)}]$. This step is performed for $j = 0, 1, \ldots, 9$ to get 10 matrices $\left\{\boldsymbol{U}_q^{(j)}, j = 0,1,\ldots,9\right\}$.

Step 3 Use (E1.4) to calculate principal components $\left\{\boldsymbol{f}_j, j = 0,1,\ldots,9\right\}$.

Step 4 Use (E1.5) to calculate $\left\{\hat{\boldsymbol{x}}_j, j = 0,1,\ldots,9\right\}$.

Step 5 Use (E1.6) to calculate $\left\{e_j, j = 0,1,\ldots,9\right\}$.

Step 6 Use (E1.7) to Identify the target class j^* for data point $\boldsymbol{x}$.

Remarks

(i) **`eigs`** is a convenient MATLAB function for implementing Step 2 of the algorithm: For a square and symmetric matric $\boldsymbol{C}$ and an integer $q > 0$, **`[Uq,Sq] = eigs(C,q)`** returns two items, namely **`Uq`** and **`Sq`**, where **`Uq`** contains q eigenvectors of $\mathbf{C}$ corresponding to the q largest eigenvalues.

(ii) It is important to realize that the 10 pairs $\{\boldsymbol{\mu}_j, \boldsymbol{U}_q^{(j)} \text{ for } j = 0,1,\ldots,9\}$ can be calculated *before* the real-time classification of new digits begins. Moreover, these pre-calculated items can be used not only for one single test sample but for the *entire* set of new digits.

4. Procedure

4.1 From the course website download the data and MATLAB function listed below.

- **`X1600.mat`** – contains a total of 16,000 handwritten digits selected at random from the training data of MNIST database. Note that **`X1600.mat`** is a matrix of size 784 × 16000, with each column representing a digit, and it contains 10 classes of digits, each class contains 1600 samples. **`X1600`** is structured so that its first 1600 columns contain 1600 digit "0", and the next 1600 columns contain 1600 digit "1", and so on.

- **`Te28.mat`** – contains 10,000 handwritten digits from testing purposes. **`Te28.mat`** is matrix of size 784 × 10,000, with each column representing a digit for testing the performance of the classifier you are to build.

- **`Lte28.mat`** – contains 10,000 integers between 0 and 9, each integer is the label of the corresponding digit from **`Te28.mat`**.

4.2 Prepare MATLAB code to implement the algorithm described in Sec. 3 above. It is desirable to prepare the code so that it can handle any number of test samples (see Remark (ii) above). It is recommended that the number q of principal axes be set to $q = 29$. Run the code with **`Te28`** as the test data.

4.3 Evaluate and report the performance of the PCA-based classifier by comparing the classification results with the labels in **`Lte28.mat`**. Report the total number of misclassifications as well as overall error rate in percentage.

4.4 Evaluate and report the efficiency of the PCA-classifier in terms of the CPU time (in seconds) consumed by the classifier per 1000 digits. Note that for a fair evaluation the CPU time for performing Steps 1 and 2 of the algorithm should *not* be counted. Instead, the CPU time should cover the time consumed by Steps 3-6.

Include your MATLAB code in the lab report please.