Assignment 5

Machine Learning, Summer term 2013, Ulrike von Luxburg

Solutions due by May 13

Linear Programming (LP): A linear program is an optimization problem with linear objective and linear constraints. For example, consider the following linear program

$$\begin{aligned} \text{minimize} \quad & 4x_1 + 3x_2 - x_3 \\ \text{subject to} \quad & -x_1 + x_2 \leq 1 \\ & 4x_1 - 2x_2 + 3x_3 \leq -2 \\ & -2x_2 - 3x_3 + 4 \leq 0 \\ \text{and} \quad & x_i \leq 0 \;;\; i = 1, 2, 3 \end{aligned}$$

We can rewrite this linear program in the standard form

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \\ \text{and} \quad & \mathbf{x} \leq \mathbf{0} \end{aligned}$$

where $x \in R^3$, $A = \begin{bmatrix} -1 & 1 & 0 \\ 4 & -2 & 3 \\ 0 & -2 & -3 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ -2 \\ -4 \end{bmatrix}$ and $c = \begin{bmatrix} 4 \\ 3 \\ -1 \end{bmatrix}$.

Exercise 1 (LP in standard form, 2 points) Write the following linear program in the standard form by determining **A,b,c** (You need to change the variable $x_3' = -x_3$).

$$\begin{aligned} \text{minimize} \quad & x_1 - 2x_2 + 4x_3 \\ \text{subject to} \quad & -x_1 + x_2 \geq 1 \\ & 3x_1 + 2x_3 \leq -1 \\ & -2x_1 - 5x_3 + 4 \leq 0 \\ \text{and} \quad & x_1, x_2 \leq 0 \\ & x_3 \geq 0 \end{aligned}$$

Exercise 2 (LP and its dual, 4 points) Find the Lagrangian dual problem for the general standard linear program. First form the Lagrangian

$$L(x, \lambda_1, \lambda_2) = \mathbf{c}^{\mathrm{T}}\mathbf{x} + \boldsymbol{\lambda_1}^{\mathrm{T}}(A\mathbf{x} - \mathbf{b}) + \boldsymbol{\lambda_2}^{\mathrm{T}}\mathbf{x}$$

where $\boldsymbol{\lambda_1}$ and $\boldsymbol{\lambda_2}$ are vectors of Lagrange multipliers.

1. Find the derivative of the Lagrangian with respect to the vector $\mathbf{x}$ and set it to zero ($\nabla_x L(x, \lambda_1, \lambda_2) = 0$). Solve it with respect to $\boldsymbol{\lambda_2}$.
2. Replace $\boldsymbol{\lambda_2}$ from part 1 in $L(x, \lambda_1, \lambda_2)$ and simplify it. The Lagrange dual problem should be in the form

$$\begin{aligned} \text{maximize} \quad & L(x, \lambda_1, \lambda_2) \\ \text{subject to} \quad & \boldsymbol{\lambda_1} \geq \mathbf{0} \\ & \boldsymbol{\lambda_2} \geq \mathbf{0} \end{aligned}$$

   with values for $L(x, \lambda_1, \lambda_2)$ and λ_2 replaced in. This is called the Lagrange dual or simply the dual problem.

Web page: `http://www.informatik.uni-hamburg.de/ML/contents/people/luxburg/teaching/2013-ss-vorlesung-ml/`
Login: "machine", Password: "learning"

3. Show that the dual of the dual linear program is the original(primal) linear program.

Exercise 3 (Dual of hard margin SVM, 5 points) Derive the dual problem of the hard margin SVM. Write all intermediate steps

$$\min_{\mathbf{w},b} \quad \frac{1}{2}\|\mathbf{w}\|^2$$
$$y_i(\mathbf{w}^{\mathrm{T}}\mathbf{x_i} - b) \geq 1 \quad i = 1, \dots, m$$

Exercise 4 (2 points) Show that, if the 1 on the right-hand side of the constraint in the hard margin SVM is replaced by some arbitrary constant $\gamma > 0$, the solution for the maximum margin hyperplane is unchanged.

Read `prepare05.pdf` for an introduction to CVX optimization package in MATLAB. An implementation of the soft margin SVM with CVX is provided there.

Exercise 5 (Cancer detection, 8 points) In this exercise you would learn a SVM that classifies cancers as either benign (-1) or malignant (+1) depending on the characteristics of sample biopsies. Load the patients data from `cancer-data.mat`. For every patient, 9 attributes are measured:
1- Clump thickness 2- Uniformity of cell size 3- Uniformity of cell shape 4- Marginal Adhesion 5- Single epithelial cell size 6- Bare nuclei 7- Bland chomatin 8- Normal nucleoli 9- Mitoses.
Train a SVM classifier on the training data. Use 10-fold cross validation to select the parameter $C \in \{0.01, 0.1, 1, 10, 50\}$. Report the train and test error. What is the effect of choosing large C on the training error?