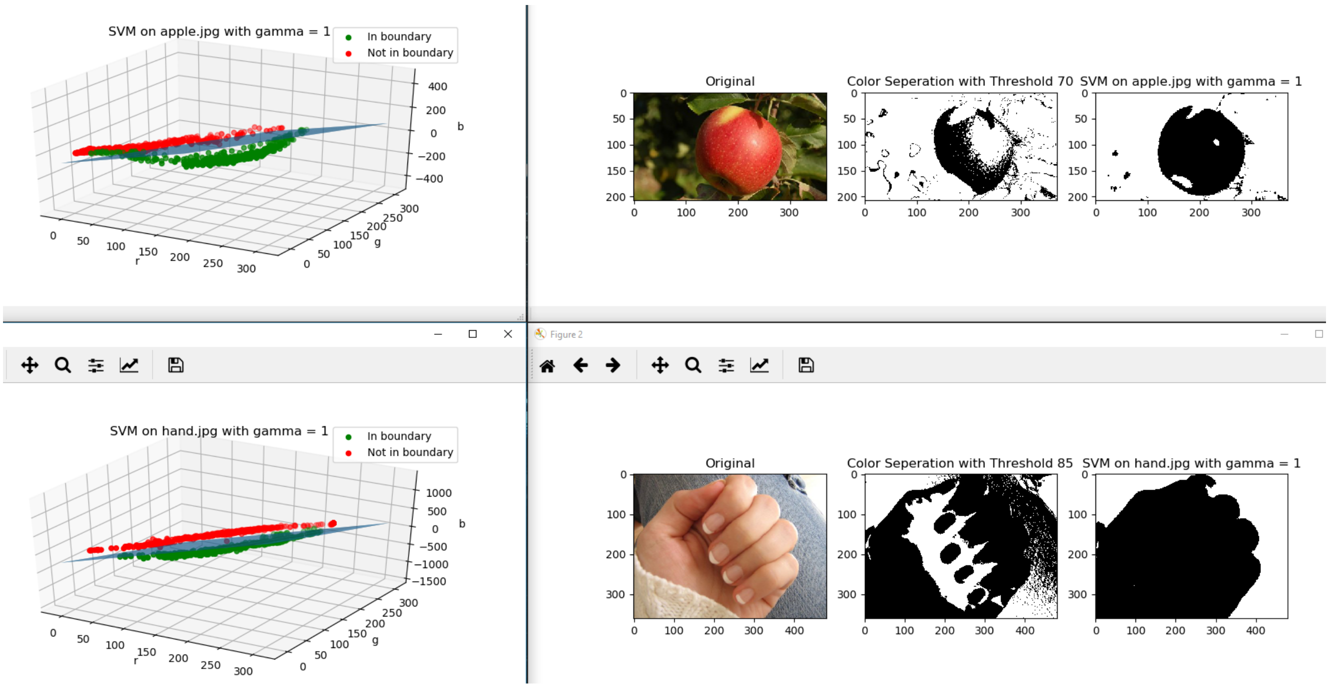
Convex Optimization Homework 5

Huafeng Fan

April 9, 2019

1. (SVM Project) The SVM project was implemented in Python 3.6.5 using cvxpy. The convex optimization formulation of the SVM problem was taken from Chapter 8 of the textbook. The training data of the images were handpicked segments of the image that were strictly inside and outside of the boundary. Then, to speed up the training stage, the input fed to train the model was a random sample of these points. Here are the plots obtained:



The top refers to results obtained with the apple picture, and the bottom refers to results obtained with the hand picture. The scatter plots on the left plots the r, g, b values of the training points, with green points in the boundary and red points not in the boundary. The light blue plane is the classifier learned through the SVM problem. The left half shows images of the original picture, distinctions made with color separation, and distinctions made through SVM respectively. It is clear that using SVM vastly outperforms simple image thresholding. Furthermore, it seems that the learned hyperplane mostly separates points based on rgb values rather than based on physical location in the image. Therefore, I would expect performance to be similar if the positions of the pixels weren't included in the feature vector, although I have not run experiments to verify this.

2. (Boyd Text Problem 5.5) We want to derive the dual problem of the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

Note that the Lagrangian of the objective function is:

$$L(x, \lambda, \nu) = c^T x + \lambda^T (Gx - h) + \nu^T (Ax - b)$$

Finding the dual function, we get

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in D} [c^T x + \lambda^T (Gx - h) + \nu^T (Ax - b)] \\ &= \inf_{x \in D} [(c^T + \lambda^T G + \nu^T A)x] - \lambda^T h - \nu^T b \\ &= \begin{cases} -\lambda^T h - \nu^T b & \text{if } c^T + \lambda^T G + \nu^T A \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Finally, we can formulate the dual problem as follows:

$$\begin{array}{ll} \text{maximize} & -\lambda^T h - \nu^T h \\ \text{subject to} & c^T + \lambda^T G + \nu^T A \succeq 0 \\ & \lambda \succeq 0 \end{array}$$

3. (Boyd Text Problem 5.12) We want to derive the dual problem of the LP

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^m \log y_i \\ \text{subject to} & y_i + a_i^T x - b_i = 0 (\forall i = 1, \ldots, m) \end{array}$$

Note that the Lagrangian of the objective function is:

$$L(x, \lambda, \nu) = -\sum_{i=1}^m \log y_i + \sum_{i=1}^m \nu_i (y_i + a_i^T x - b_i)$$

Finding the dual function, we get

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x,y} [-\sum_{i=1}^m \log y_i + \sum_{i=1}^m \nu_i (y_i + a_i^T x - b_i)] \\ &= \inf_{x,y} [-\sum_{i=1}^m \log y_i + \nu^T y - \nu^T b + \nu^T A x] \end{aligned}$$

Note that $\nu^T A x$ can grow arbitrarily small with x unless $\nu^T A = A^T \nu = 0$. Also, note that it's not clear which value of y will minimize this expression. To resolve this, we want to find the y that minimizes:

$$t(y) = \sum_{i=1}^m (\nu_i y_i - \log y_i)$$

To do this, find the derivative and set it equal to zero.

$$t'(y) = \sum_{i=1}^m (\nu_i - \frac{1}{y_i}) = 0 \implies y_i = \frac{1}{\nu_i} (\forall i)$$

so the value of y will minimize $L(x, \lambda, \nu)$ is that in which $y_i = \frac{1}{\nu_i} (\forall i)$. Therefore, we have that

$$g(\lambda, \nu) = \begin{cases} \sum_{i=1}^m \log \nu_i + m - \nu^T b & \text{if } A^T \nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Finally, we can formulate the dual problem as follows:

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^m \log \nu_i + m - \nu^T b \\ \text{subject to} & A^T \nu = 0 \end{array}$$