Math189R SP19
Homework 5
Monday, Mar 11, 2019

Feel free to work with other students, but make sure you write up the homework and code on your own (no copying homework *or* code; no pair programming). Feel free to ask students or instructors for help debugging code or whatever else, though.

1 (Murphy 12.5 - Deriving the Residual Error for PCA) It may be helpful to reference section 12.2.2 of Murphy.

(a) Prove that

$$\left\| \mathbf{x}_i - \sum_{j=1}^{k} z_{ij}\mathbf{v}_j \right\|^2 = \mathbf{x}_i^\top \mathbf{x}_i - \sum_{j=1}^{k} \mathbf{v}_j^\top \mathbf{x}_i \mathbf{x}_i^\top \mathbf{v}_j.$$

Hint: first consider the case when $k = 2$. Use the fact that $\mathbf{v}_i^\top \mathbf{v}_j$ is 1 if $i = j$ and 0 otherwise. Recall that $z_{ij} = \mathbf{x}_i^\top \mathbf{v}_j$.

(b) Now show that

$$J_k = \frac{1}{n} \sum_{i=1}^{n} \left(\mathbf{x}_i^\top \mathbf{x}_i - \sum_{j=1}^{k} \mathbf{v}_j^\top \mathbf{x}_i \mathbf{x}_i^\top \mathbf{v}_j \right) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i^\top \mathbf{x}_i - \sum_{j=1}^{k} \lambda_j.$$

Hint: recall that $\mathbf{v}_j^\top \boldsymbol{\Sigma} \mathbf{v}_j = \lambda_j \mathbf{v}_j^\top \mathbf{v}_j = \lambda_j$.

(c) If $k = d$ there is no truncation, so $J_d = 0$. Use this to show that the error from only using $k < d$ terms is given by

$$J_k = \sum_{j=k+1}^{d} \lambda_j.$$

Hint: partition the sum $\sum_{j=1}^{d} \lambda_j$ into $\sum_{j=1}^{k} \lambda_j$ and $\sum_{j=k+1}^{d} \lambda_j$.

a) I had to look at the solution for this one. Again, I'll summarize:

$$\left\| \mathbf{x}_i - \sum_{j=1}^{k} z_{ij}\mathbf{v}_j \right\|^2 = (\mathbf{x}_i - \sum_{j=1}^{k} z_{ij}\mathbf{v}_j)^T (\mathbf{x}_i - \sum_{j=1}^{k} z_{ij}\mathbf{v}_j)$$

Simplify by bringing $\mathbf{x_i}^T$ into the sum, and since $\mathbf{v_i}^T \mathbf{v_j} = 1$ if and only if $i = j$, we finally get:

$$\mathbf{x_i}^T \mathbf{x_i} - \sum_{j=1}^{k} \mathbf{v_j}^T \mathbf{x_i} \mathbf{x_i}^T \mathbf{v_j}$$

as desired

b)

$$J_k = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x_i}^T\mathbf{x_i} - \sum_{j=1}^{K}\mathbf{v_j}^T\mathbf{x_i}\mathbf{x_i}^T\mathbf{v_j})$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathbf{x_i}^T\mathbf{x_i} - \sum_{j=1}^{k}\mathbf{v}_j^T\Sigma\mathbf{v}_j$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathbf{x_i}^T\mathbf{x_i} - \sum_{j=1}^{k}\lambda_j$$

c) Looked at the solutions for reference for this one. We know that:

$$J_k = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i^T\mathbf{x}_i - \sum_{j=1}^{d}\lambda_j + \sum_{j=k+1}^{d}\lambda_j = \sum_{j=k+1}^{d}\lambda_j$$

As Prof Gu mentioned during lecture, the sum of the unused eigenvalues is equal to the reconstruction error of a PCA projection.

∎

2 (ℓ_1-Regularization) Consider the ℓ_1 norm of a vector $\mathbf{x} \in \mathbb{R}^n$:

$$\|\mathbf{x}\|_1 = \sum_i |\mathbf{x}_i|.$$

Draw the norm-ball $B_k = \{\mathbf{x} : \|\mathbf{x}\|_1 \leq k\}$ for $k = 1$. On the same graph, draw the Euclidean norm-ball $A_k = \{\mathbf{x} : \|\mathbf{x}\|_2 \leq k\}$ for $k = 1$ behind the first plot. (Do not need to write any code, draw the graph by hand).

Show that the optimization problem

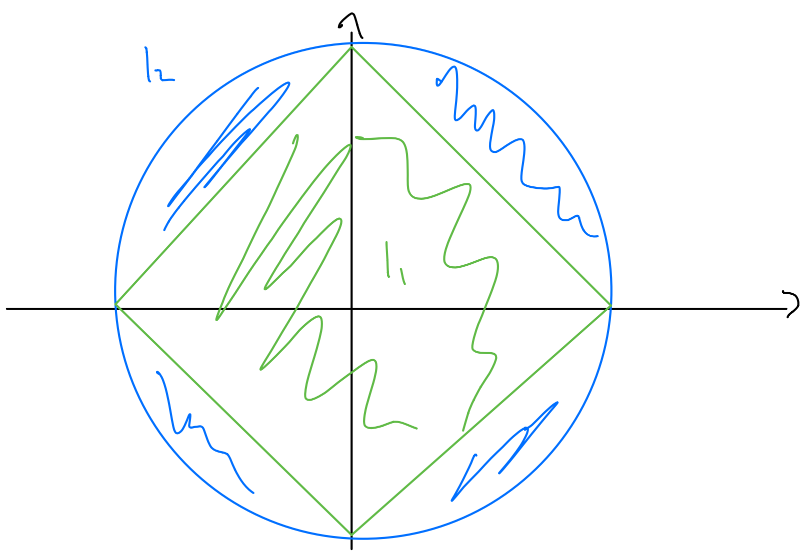
$$\begin{aligned} \text{minimize: } & f(\mathbf{x}) \\ \text{subj. to: } & \|\mathbf{x}\|_p \leq k \end{aligned}$$

is equivalent to

$$\text{minimize: } f(\mathbf{x}) + \lambda \|\mathbf{x}\|_p$$

(hint: create the Lagrangian). With this knowledge, and the plots given above, argue why using ℓ_1 regularization (adding a $\lambda\|\mathbf{x}\|_1$ term to the objective) will give sparser solutions than using ℓ_2 regularization for suitably large λ.

Graph:



Per the class notes,

$$infsupL(\mathbf{x}, \lambda) = infsupf(\mathbf{x}) + \lambda(||\mathbf{x}||_p - k)$$

Now, referring to the solution guide, we can flip the inf and the sup because it is in its dual, such that:

$$infsupf(\mathbf{x}) + \lambda(||\mathbf{x}||_p - k) = supg(\lambda)$$

By optimizing x, we will solve the minimization of $f(\mathbf{x}) + \lambda||\mathbf{x}||_p$ for some value of$\lambda >= 0$. After looking at the solution, I still don't quite understand what is going on. I will be sure to ask Prof Gu either during class or I will set up an appointment with her. ∎