

Homework 2

Please answer the following questions in complete sentences in a typed manuscript and submit the solution to me in class on January 31th, 2012.

In this homework, we probe a bit of theory around the topics we've covered in class so far. The goal is not to provide tedious work, but useful results that'll rely upon later. For instance, the first problem tackles subspaces and null-spaces, which will frequently arise in our discussion of constrained optimization. The second problem asks you use the tools in class to discuss a variation on least squares that often arises in many applications. The third problem deals with convexity of a subspace in relationship to the constrained least squares problem. Finally, the fourth problem shows a different way of looking at a least squares problem.

Problem 1: Constrained least squares theory

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times n}$. Consider the constrained least squares problem:

$$\begin{aligned} & \text{minimize} && \|\mathbf{b} - \mathbf{A}\mathbf{x}\| \\ & \text{subject to} && \mathbf{B}\mathbf{x} = \mathbf{d} \end{aligned} .$$

1. Show that $\text{rank}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) = n$ is equivalent to the condition that the intersection between the null-spaces of A and B only has the trivial vector $\mathbf{x} = 0$.
2. Show that the problem may not have any solution if $\text{rank}(\mathbf{B}) < p$.
3. Show that the problem always has a unique solution, for any \mathbf{d} , when $\text{rank}(\mathbf{B}) = p$ and $\text{rank}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) = n$.
4. Show that if the problem has a unique solution, for any \mathbf{d} , then $\text{rank}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) = n$. (You can combine this with the proof of the last one if you wish.)

Problem 2: Ridge Regression

The Ridge Regression problem is a variant of least squares:

$$\text{minimize} \quad \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_2^2.$$

This is also known as Tikhonov regularization.

1. Show that this problem always has a unique solution, for *any* \mathbf{A} if $\lambda > 0$, using the theory discussed in class so far.

This property is one aspect of the reason that Ridge Regression is used. It is also a common regularization method that can help avoid overfitting in a regression problem.

1. Use the SVD of \mathbf{A} to characterize the solution as a function of λ .
2. What is the solution when $\lambda \rightarrow \infty$?
What is the solution when $\lambda \rightarrow 0$?

Problem 3: Convexity

1. Show that $f(x) = \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$ is a convex function. Feel free to use the result proved on the last homework.
2. Show that the null-space of a matrix is a convex set.
3. Explain why these results allow us to conclude that constrained least squares is a convex optimization problem. (See Wikipedia for the definition of a convex problem.)

Problem 4: Alternate formulations of Least Squares

Consider the constrained least squares problem:

$$\begin{array}{ll} \text{minimize}_{\mathbf{r}, \mathbf{y}} & \|\mathbf{r}\|_2 \\ \text{subject to} & \mathbf{r} = \mathbf{b} - \mathbf{C}\mathbf{y} \end{array}$$

where $\mathbf{C} \in \mathbb{R}^{m \times n}$, $n \leq m$ and rank n .

1. Convert this problem into into the standard constrained least squares form.
2. Form the augmented system from the Lagrangian as we did in class.
3. Manipulate this problem to arrive at the normal equations for a least-squares problem: $\mathbf{C}^T \mathbf{C} \mathbf{y} = \mathbf{C}^T \mathbf{b}$.
Discuss any advantages of the systems at intermediate steps.