AM 160 - SciML:

Dante Buhl

January 23, 2025

Problem 1:

Subproblem 1.1:

Show that $\Theta = A^\dagger y$ is the solution such that $||\Theta||_2$ is the minimum out of all the infinite solutions to $y = A\Theta$. Consider $A \in \mathbb{R}^{n\times p}$ where $p \gg n$.

Proof. We begin by writing the SVD of A which is part of the construction of the psuedoinverse $A^\dagger$.

$$A = U\Sigma V^*$$
$$A^\dagger = V\Sigma^{-1}U^*$$

where U and V are unitary matrices, and Σ^{-1} is the transpose of Σ containing the reciprocal of each singular value (in order) along the diagonal, i.e.

$$\Sigma^{-1} = \left[\begin{array}{ccc} 1/\sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1/\sigma_n \\ \hline & \mathbf{0} & \end{array}\right]$$

Next we need to understand why Θ in this context has infinitely many solutions. Let us assume that A is rank n. The implication is that there are at most n basis vectors in $\mathbb{R}^p$ which are not in the null space of the transformation A. We can write Θ as a linear combination of basis vectors which span $\mathbb{R}^p$

$$\Theta = c_1v_1 + \ldots + c_nv_n + \ldots + c_pv_p$$

Notice though, however, that only n of these vectors are in the kernel of A (and let us assume it is the first n vectors for convenience). We have then that all vectors $v_{n+1}, \ldots, v_p$ in the linear combination of Θ do not affect the solution. Therefore, an infinite number of solutions Θ can be created by adding the linear combination of $v_{n+1}, \ldots, v_p$ to any solution of $y = A\Theta$.

In order to see why the psuedouinverse $A^\dagger$ yields the minimum solution is because it projects the p dimension problem into a n dimension problem, i.e.

$$y = A\Theta$$

$$y = U\Sigma V^*\Theta$$

$$y = UC$$

where C is an $n \times 1$ vector which literally contains the coefficients of the linear combination for the first n basis vectors of Θ scaled by their corresponding singular value σ_i, i.e. $C_i = c_i\sigma_i$.

Finally, we solve for C (which has one unique solution since U is full rank) using the inverse of U which exists since U is unitary.

$$C = U^*y$$

Notice that this constructs Θ out of the minimum number of basis vectors in order to span $\mathbb{R}^n$ that is, for any y given. Then looking at the 2-norm of Θ we have,

$$\begin{aligned}
||\Theta||_2 &= ||c_1v_1||_2 + \ldots + ||c_nv_n||_2 \\
&= |c_1|||v_1||_2 + \ldots + |c_n|||v_n||_2 \\
&= |c_1| + \ldots + |c_n|
\end{aligned}$$

where we can decompose the 2-norm in this way because the basis vectors v_i are orthogonal to each other in the 2-norm. Notice that the addition of any additional $\mathbb{R}^p$ basis vectors will only increase the 2-norm of Θ. We conclude then that solving for Θ using the psuedoinverse of A yields Θ such that the 2-norm of Θ is the minimum out of all possible Θ which solve $y = A\Theta$.

The proof at the end of Lecture Notes 4, essentially demonstrates this as well. The basis of the proof is that any of the infinitely many solutions can be constructed as

$$\Theta = \Theta' + \Theta^*$$

and proceeds to show that $\langle \Theta', \Theta^* \rangle = 0$, i.e. that Θ' is orthogonal to Θ^*. As I have shown earlier, the infinite solutions for Θ arise by introducing a linear combination of the vectors in the Null space of A, i.e.

$$\Theta = \Theta^* + c_{n+1}v_{n+1} + \ldots + c_p v_p$$

$$\Theta' = c_{n+1}v_{n+1} + \ldots + c_p v_p$$

Finally, since Θ' is constructed of basis vectors of $\mathbb{R}^p$ all of which are orthogonal to Θ^*, we have that $\langle \Theta', \Theta^* \rangle = 0$. The rest of the proof in the lecture notes holds (assuming $\Theta' \neq 0$),

$$\begin{aligned}
\langle \Theta, \Theta \rangle &= \langle \Theta' + \Theta^*, \Theta' + \Theta^* \rangle \\
&= \Theta'^T\Theta' + \Theta'^T\Theta^* + \Theta^{*T}\Theta' + \Theta^{*T}\Theta^* \\
&= \langle \Theta', \Theta' \rangle + 2\langle \Theta', \Theta^* \rangle + \langle \Theta^*, \Theta^* \rangle \\
&= \langle \Theta', \Theta' \rangle + \langle \Theta^*, \Theta^* \rangle > \langle \Theta^*, \Theta^* \rangle
\end{aligned}$$

where both $\langle \Theta', \Theta' \rangle$ and $\langle \Theta^*, \Theta^* \rangle$ are semi-positive definite terms. Therefore, we have that $\Theta^* = \text{argmin}_{y=A\Theta}||\Theta||_2$.

$\square$

Subproblem 1.2:

In order to demonstrate double descent using a random Fourier feature matrix A, we perform a linera regression for the function,

$$y = 2x + \cos(25x) + r \tag{1}$$

where r is gaussian noise sampled from the standard normal distribution $\mathcal{N}(0, 1)$. Since the purpose of the linear regression

$$y = A\Theta$$

is to build a fourier cosine series which interpolates (1), where the vector Θ is the coefficient vector for each random fourier mode $\phi_i(x) = \cos(2\pi w_i x)$, where $w_i \sim \mathcal{N}(0, 1)$. For my work specifically, I have chosen to use 50 data points for each of the training and the test set on the domain $x \in [1, 10]$. It is necesary for both data sets to be taken from the same domain, as fourier series usually only interpolate functions in a finite domain.

As a result, the number of features used in this regression problem will be start at 5 features ($p \ll n$), and

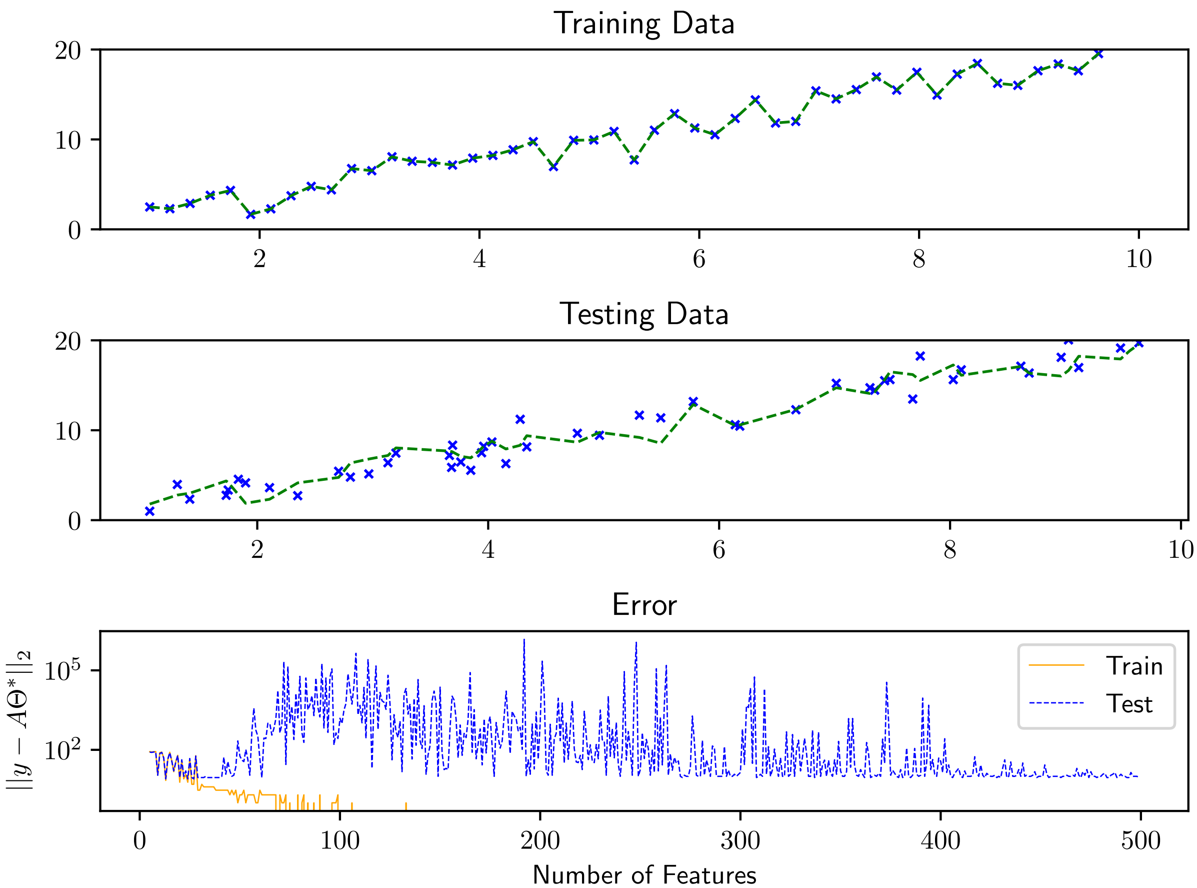

Figure 1: Training, Test, and Error data from the linear regression posed in subproblem 2. The training and test data plots show how well the linear regression fits to the data given in both the training and testing cases for $p = 500$. The error plot shows both the training and testing error as a function of the number of features p, in order to demonstrate double descent.

end at 500 features ($p \gg n$) with increments of 1. Notice that depending on the type of system being solved (overdetermined v.s. underdetermined) the method of obtaining Θ will change from a least squares solution to a psuedoinverse solution, i.e.

$$\Theta = \text{argmin}_{y=A\Theta}||y - A\Theta||_2, \quad p < n$$

$$\Theta = A^\dagger y, \quad p \geq n$$

Thus, we obtain the error v.s. feature plot that is seen in figure 1, where there appears to be a double descent which settles to the minimum test error near $p \approx 400$. Note that the minimum test error seems to settle at 10^1 which seems natural due to the noise in the data sets which is of $O(10^0)$, and then scaled by the number of test points which is of order $O(10^1)$, hence obtaining an average error for the whole data set of $O(10^1)$.