M442, Spring 2022 Practice Problems for the Midterm

The midterm exam for M442 will be Wednesday, March 9, 7-9 p.m., in Blocker 150. The exam will cover the following topics: linear and nonlinear regression, including theory and implementation; dimensional analysis, including theory and implementation (both with regression and structured experiments); and modeling biological processes with ODE, including single-species models, models with interacting species, and epidemic models.

The exam will consist of two parts: Part 1 will not require MATLAB, while Part 2 will require MATLAB. Students will not be allowed to use notes or references for Part 1, and students must submit Part 1 either in person or via Canvas before beginning Part 2. For Part 2, students can access any notes, references, M-files etc., except no communication between students is allowed.

For some questions, students may be expected to access data files from the course web site.

The problems in homework assignments 2, 3, and 4 serve as good practice for the exam, and solutions can be found in Canvas under the *files* link. The problems below are intended to provide students with some additional practice, especially on topics not covered in the homework. These problems are not assigned to be turned in, and solutions are included.

Practice Problems

1. Suppose we have data $\{(x_k, y_k)\}_{k=1}^N$ for which x_k is the same value for each $k \in \{1, 2, \dots, N\}$. I.e., suppose we have $x_k = \mu_x$ for each $k \in \{1, 2, \dots, N\}$. For a line fit $y = p_1 + p_2 x$, write down the normal equation for this data and characterize its solutions.

2. The formation of a black hole in space occurs when the mass m within a sphere exceeds a threshold that depends on the radius R of the sphere, the speed of light in vacuum c, and Newton's gravitational constant G. Use dimensional analysis to find a general form for this threshold.

Note. $[G] = M^{-1}L^3T^{-2}$.

3. Use the method of dimensional analysis to determine a general form for the period P of a pendulum of length l released from rest at angle θ. Ignore air resistance, and note that the angle θ between the pendulum and the vertical is dimensionless.

4. Continuing with Problem 3, pendulum data from our reference on dimensional analysis by Palmer is given in the M-file *pdata.m* (available on the course web site). Use this data, and your results from Problem 3, to obtain a model for pendulum motion as a function of l, g, and θ. In this case, it's known from Newtonian mechanics that the exact formula for P is

$$P = 4\sqrt{\frac{l}{g}}K(\sin\frac{\theta}{2}),$$

where K is the *complete elliptic integral of the first kind*

$$K(x) := \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{1 - x^2\sin^2\phi}}.$$

For the fixed values $l = 1$ and $g = 9.81$, give a graphical comparison of your model values of P along with the (theoretically) exact values for $\theta \in [0, \frac{\pi}{2}]$.
Note. MATLAB's built-in function for K is *ellipke.m,* with one adjustment:

$$K(x) = \text{ellipke}(x^2).$$

5. For a fluid such as oil moving through a pipe, the velocity v at a certain point along the pipe will generally depend on the diameter D of the pipe, the density ρ of the fluid, the viscosity μ of the fluid, and the pressure drop $\frac{dp}{dx}$ at the point.
a. Find a general form for v as a function of v, D, ρ, μ, and $\frac{dp}{dx}$.
b. Suppose you would like to build a large pipeline with $D = 10$ m, but you would first like to run experiments on a table model with $D = .05$ m. How must the values of viscosity, density, and fluid drop for the structured experiment be related to those of the original pipeline?
6. When a raccoon is infected by the rabies virus, one of two (equally likely) things can happen:

1. The raccoon develops *furious rabies*, in which case he becomes hyperactive and is quick to attack other raccoons.

2. The raccoon develops *dumb rabies,* in which case he becomes paralyzed and does not spread the disease.

In either case, the raccoon will not recover, and will die within a week of becoming infected. Develop a model for the spread of the rabies virus through an isolated population of raccoons. Which parameter in your model is determined by the fact that infected raccoons will die within a week of becoming infected, and what is the value of this parameter?
7. The West Nile Virus is carried predominately by birds and mosquitoes: When an infected mosquito bites a susceptible bird, the virus remains in the bird's blood for two or three days, long enough to infect a great number of mosquitoes. Assume that when a bird recovers it becomes susceptible again. That is, it neither dies nor becomes immune. Taking into account four variables—number of uninfected mosquitoes, $m(t)$, number of infected mosquitoes, $n(t)$, number of uninfected birds, $b(t)$, and number of infected birds, $c(t)$—develop a model for the spread of West Nile Virus through a population of birds and mosquitoes. While you should assume mosquitoes are born and die, since the time horizon researchers are primarily concerned with is, say, one summer, you should consider the total bird population $(b(t)+c(t))$ to be fixed. Determine which parameter value in your model corresponds with the virus remaining in the bird's blood for two or three days, and approximate the value of this parameter.

Solutions

1. In this case, the design matrix is

$$F = \begin{pmatrix} 1 & \mu_x \\ 1 & \mu_x \\ \vdots & \vdots \\ 1 & \mu_x \end{pmatrix},$$

so

$$F^T F = \begin{pmatrix} N & N\mu_x \\ N\mu_x & N\mu_x^2 \end{pmatrix} \quad \text{and} \quad F^T \vec{y} = \begin{pmatrix} \sum_{k=1}^N y_k \\ \mu_x \sum_{k=1}^N y_k \end{pmatrix}.$$

We see that $F^T F \vec{p} = F^T \vec{y}$ reduces to a single equation

$$p_1 + \mu_x p_2 = \frac{1}{N} \sum_{k=1}^N y_k = \mu_y.$$

In this way, we see that the normal equation has an infinite number of solutions; in particular any pair (p_1, p_2) satisfying

$$p_1 = \mu_x p_2 + \mu_y.$$

2. We look for a threshold of the form

$$m = kR^a c^b G^d,$$

with corresponding dimensions

$$M = L^a L^b T^{-b} M^{-d} L^{3d} T^{-2d}.$$

The dimensions equations are seen to be

$$\begin{aligned} L:&\ 0 = a + b + 3d \\ M:&\ 1 = -d \\ T:&\ 0 = -b - 2d. \end{aligned}$$

We see immediately that $d = -1$, $b = 2$, and $a = 1$, and conclude that the relation we're looking for is

$$m = k\frac{c^2 R}{G}.$$

Note. The precise value of k can be determined (by other methods) to be $k = \frac{1}{2}$.

3. In the absence of air resistance, the period of a pendulum should depend on the length of the pendulum, l, the force of gravity, g, and the angle the pendulum is pulled from the vertical, θ. We look for dimensionless products

$$\pi = l^a \theta^b g^c P^d,$$

with dimensions

$$1 = L^a L^c T^{-2c} T^d.$$

This gives the dimension equations

$$\begin{aligned} L&: 0 = a + c \\ T&: 0 = -2c + d. \end{aligned}$$

For π_1 we choose

$$\pi_1 = \theta,$$

while for π_2 we choose $d = 1$ to get $c = \frac{1}{2}$ and $a = -\frac{1}{2}$, giving

$$\pi_2 = P\sqrt{\frac{g}{l}}.$$

According to Buckingham's Theorem we can express the equation for the period of a pendulum in the form

$$f(\pi_1, \pi_2) = 0,$$

and the Implicit Function Theorem suggests there exists some function $\phi(\pi_1)$ so that

$$\pi_2 = \phi(\pi_1).$$

I.e.,

$$P\sqrt{\frac{g}{l}} = \phi(\theta),$$

and so

$$P = \phi(\theta)\sqrt{\frac{l}{g}}.$$

4. We solve Problem 4 with the M-file *pdatafit.m*.

```
%PDATAFIT: MATLAB script M-file that uses the pendulum data in pdata.m
%to find a function for the period of a pendulum in terms of
%length l, angle theta, and gravity g
%
%Define data
pdata;
p=polyfit(theta,P.*sqrt(g./l),1)
plot(theta,P.*sqrt(g./l),'o',theta,p(1)*theta+p(2))
title({'Plot of {\pi}_2 as a function of {\pi}_1 = \theta', 'Along with best-fit
line'},'FontSize',15)
axis equal
%Comparison with exact solution
clear l;
l = 1;
theta = linspace(0,pi/2,100);
Pexact = 4*sqrt(l/g).*ellipke(sin(theta/2).^2);
Pmodel = sqrt(l/g)*(p(1)*theta+p(2));
figure
plot(theta,Pexact,theta,Pmodel,'--')
title({'Plot of Model Periods Along with Exact Periods', 'With l = 1, 0 <=
{\theta} <= {\pi}/2'},'FontSize',15)
legend('exact periods','model periods')
```

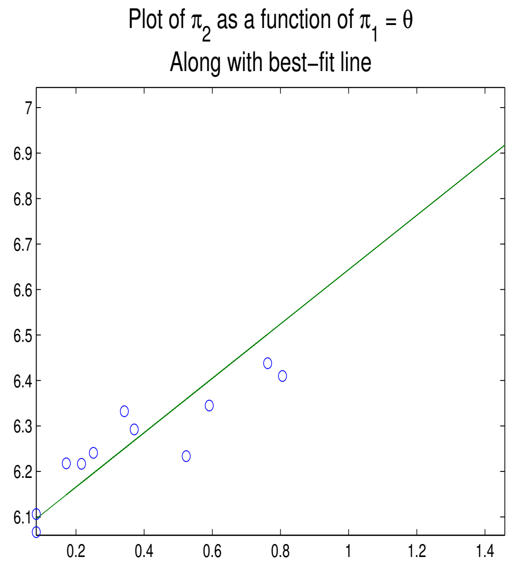

Figure 1: Fit of $\pi_2 = P\sqrt{\frac{g}{l}}$ as a function of $\pi_1 = \theta$.

The fit for $\pi_2 = \phi(\pi_1)$ is shown in Figure 1.
We find $p_1 = .5976$ and $p_2 = 6.0462$, so that our model is

$$P(l, g, \theta) = \sqrt{\frac{l}{g}}(.5976\theta + 6.0462).$$

The comparision of this model with the exact solution is shown in Figure 2.
5. For (a), we must use dimensionless products, so we write

$$\pi = \pi(D, \rho, \frac{dp}{dx}, \mu, v) = D^a \rho^b (\frac{dp}{dx})^c \mu^d v^e,$$

with dimensions

$$1 = L^a M^b L^{-3b} M^c L^{-2c} T^{-2c} M^d L^{-d} T^{-d} L^e T^{-e}.$$

The dimension equations are

$$\begin{aligned} L &: 0 = a - 3b - 2c - d + e \\ M &: 0 = b + c + d \\ T &: 0 = -2c - d - e. \end{aligned}$$

In this case, we have two degrees of freedom. For π_1, we set $e = 0$ and $d = 1$ to get the reduced system

$$\begin{aligned} 0 &= a - 3b - 2c - 1 \\ 0 &= b + c + 1 \\ 0 &= -2c - 1, \end{aligned}$$

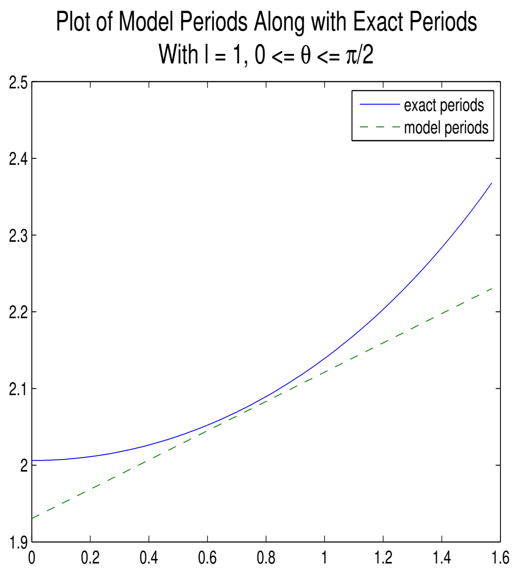

Figure 2: Comparison of model periods with theoretically exact periods.

from which we have $c = -\frac{1}{2}$, $b = -\frac{1}{2}$, and $a = -\frac{3}{2}$. We conclude

$$\pi_1 = \frac{\mu}{\sqrt{\rho(\frac{dp}{dx})D^3}}.$$

For π_2, we take $e = 1$ and in this case $c = 0$ (not necessary, but gives Reynold's number) to get the reduced system

$$\begin{aligned} 0 &= a - 3b - d + 1 \\ 0 &= b + d \\ 0 &= -d - 1, \end{aligned}$$

from which we have $d = -1$, $b = 1$, and $a = 1$. We conclude

$$\pi_2 = \frac{v\rho D}{\mu},$$

which is Reynold's number. According to Buckingham's Theorem and the Implicit Function Theorem, there will generally be a function ϕ so that

$$\pi_2 = \phi(\pi_1).$$

This means

$$\frac{v\rho D}{\mu} = \phi(\frac{\mu}{\sqrt{\rho(\frac{dp}{dx})D^3}}),$$

and so

$$v = \frac{\mu}{\rho D}\phi(\frac{\mu}{\sqrt{\rho(\frac{dp}{dx})D^3}}).$$

For (b), we require

$$\frac{\mu}{\sqrt{\rho(\frac{dp}{dx})D^3}} = \frac{\mu_e}{\sqrt{\rho_e(\frac{dp}{dx})_e D_e^3}},$$

which we can rearrange as

$$\frac{\mu}{\sqrt{\rho(\frac{dp}{dx})}} = \sqrt{\frac{D^3}{D_e^3}}\frac{\mu_e}{\sqrt{\rho_e(\frac{dp}{dx})_e}}.$$

For the given values $D = 10$ and $D_e = .05$ we get the relation

$$\frac{\mu}{\sqrt{\rho(\frac{dp}{dx})}} = (2.8284 \times 10^3)\frac{\mu_e}{\sqrt{\rho_e(\frac{dp}{dx})_e}}.$$

6. We define variables as follows:

$$\begin{aligned}
y_1(t) &= \#\text{of susceptible raccoons at time } t \\
y_2(t) &= \#\text{who have furious rabies at time } t \\
y_3(t) &= \#\text{who have dumb rabies at time } t.
\end{aligned}$$

In terms of these variables, the simplest reasonable model is

$$\begin{aligned}
\frac{dy_1}{dt} &= -\, a y_1 y_2 \\
\frac{dy_2}{dt} &= \frac{a}{2} y_1 y_2 - b y_2 \\
\frac{dy_3}{dt} &= \frac{a}{2} y_1 y_2 - b y_3.
\end{aligned}$$

Clearly, we can add a removed population, but it's not required since this is a closed system. If we measure time in days,

$$b \geq \frac{1}{7}$$

(interpreting "within a week" to mean $\frac{1}{b} \leq 7$).

7. The simplest reasonable model is,

$$\begin{aligned}
\frac{dm}{dt} &= k_1 m - k_2 m c \\
\frac{dn}{dt} &= k_2 m c - k_3 n \\
\frac{db}{dt} &= -\, k_4 b n + k_5 c \\
\frac{dc}{dt} &= k_4 b n - k_5 c.
\end{aligned}$$

First equation: The term $k_1 m$ denotes the birth of mosquitoes, with birth rate faster than death rate; the term $-k_2 m c$ denotes uninfected mosquitoes infected by biting infected birds.

Second equation: The term k_2mc again denotes uninfected mosquitoes becoming infected; the term $-k_3n$ denotes the death rate of infected mosquitoes (no infected mosquitoes are born). *Third equation*: The term $-k_4bn$ denotes uninfected birds being infected by mosquitoes; the term k_5c denotes the rate at which birds quit carrying the virus (2-3 days) and consequently become uninfected. *Fourth equation*: The term k_4bn again denotes the number of birds becoming infected; $-k_5c$ denotes the loss of infected birds when they cease carrying the virus. Observe that it's clear from the last two equations that $\frac{d}{dt}(b(t)+c(t))=0$.

The parameter is k_5 (bird recovery rate), and it's value is approximately $\frac{1}{2}$ to $\frac{1}{3}$.