Contents

AA 203 HW 8 Question 1

Somrita Banerjee

```
clc
clear all
close all
rng('default') % For reproducibility
```

Part b

```
x0 = 1;
x1 = 0.5;
a_true = 1;
b_true = -0.1;
num_trials = 100;
N_values = [10, 100, 1000];
for N = N_values
    theta_estimates = zeros(2, num_trials);
    for nt=1:num_trials
        x = zeros(N+1, 1);
        for i=0:N
            if i == 0
                x(i+1) = x0;
            elseif i == 1
                x(i+1) = x1;
            else
                x(i+1) = a_true*x(i) + b_true*x(i-1) + normrnd(0,1);
            end
        end
        Y = x(3:end);
        H = [x(2:end-1) x(1:end-2)];
        theta = H\Y;
        theta_estimates(:,nt) = theta;
    end
    lse = mean(theta_estimates,2);
    stddev = std(theta_estimates,0,2);
    fprintf('N = %d : Least squares estimate of (a,b) is mean (%.2f, %.2f) with std dev (%.2f, %.2f)\n',N, lse(1),lse(2), stddev(1), stddev(2));
end
fprintf('We can see that as the amount of data used increases (i.e. N increases),\n our estimates for a and b get better. \n The mean gets closer to the true value
```

```
N = 10 : Least squares estimate of (a,b) is mean (0.92, -0.14) with std dev (0.34, 0.40)
N = 100 : Least squares estimate of (a,b) is mean (1.01, -0.13) with std dev (0.11, 0.11)
N = 1000 : Least squares estimate of (a,b) is mean (1.00, -0.10) with std dev (0.03, 0.03)
We can see that as the amount of data used increases (i.e. N increases),
 our estimates for a and b get better.
 The mean gets closer to the true value and the standard deviation decreases.
 This is in accordance with what we expect which is that more data will
 result in a better estimate.
 True values of a and b are (1.00, -0.10)
```

Part c

```
for N = N_values
    theta_estimates = zeros(2, num_trials);
    for nt=1:num_trials
        x = zeros(N+1, 1);
        noise = zeros(N+1,1);
        for i=0:N
            if i == 0
                x(i+1) = x0;
            elseif i == 1
                x(i+1) = x1;
            else
                if i == 2
                    noise(i) = normrnd(0,1);
                else
                    noise(i) = normrnd(0, sqrt(2*abs(noise(i-1))));
                end
                x(i+1) = a_true*x(i) + b_true*x(i-1) + noise(i);
            end
        end
        Y = x(3:end);
        H = [x(2:end-1) x(1:end-2)];
        theta = H\Y;
        theta_estimates(:,nt) = theta;
```

```
        end
    lse = mean(theta_estimates,2);
    stddev = std(theta_estimates,0,2);
    fprintf('N = %d : Least squares estimate of (a,b) is mean (%.2f, %.2f) with std dev (%.2f, %.2f)\n',N, lse(1),lse(2), stddev(1), stddev(2));
end
fprintf('These estimates are worse than the estimates from part b. \n The mean is further from the true values and the standard deviation is higher. \n This makes s
```

```
N = 10 : Least squares estimate of (a,b) is mean (0.92, -0.16) with std dev (0.49, 0.46)
N = 100 : Least squares estimate of (a,b) is mean (0.99, -0.10) with std dev (0.18, 0.16)
N = 1000 : Least squares estimate of (a,b) is mean (0.98, -0.09) with std dev (0.06, 0.05)
These estimates are worse than the estimates from part b.
 The mean is further from the true values and the standard deviation is higher.
 This makes sense because the noise is no longer independent between each sample,
 which adversely affects the quality of our estimate.
 True values of a and b are (1.00, -0.10)
```

Part d

```
x0 = 1;
x1 = 0.5;
c_true = 0.1;
num_trials = 100;
N_values = [10, 100, 1000];
fprintf('Sample independent noise\n')
for N = N_values
    theta_estimates = zeros(1, num_trials);
    for nt=1:num_trials
        x = zeros(N+1, 1);
        for i=0:N
            if i == 0
                x(i+1) = x0;
            elseif i == 1
                x(i+1) = x1;
            else
                x(i+1) = c_true*x(i)*x(i-1) + normrnd(0,1);
            end
        end
        Y = x(3:end);
        H = [x(2:end-1).*x(1:end-2)];
        theta = H\Y;
        theta_estimates(:,nt) = theta;
    end
    lse = mean(theta_estimates,2);
    stddev = std(theta_estimates,0,2);
    fprintf('N = %d : Least squares estimate of c is mean %.2f with std dev %.2f\n',N, lse, stddev);
end
```

```
Sample independent noise
N = 10 : Least squares estimate of c is mean 0.07 with std dev 0.44
N = 100 : Least squares estimate of c is mean 0.09 with std dev 0.09
N = 1000 : Least squares estimate of c is mean 0.10 with std dev 0.03
```