Assignment 2

02610 Optimization and Data Fitting – Anders Hørsted (s082382)

Question 1: Fitting to Air Pollution Data

In this question we are going to fit sums of sines and cosines to the Air Pollution Data given in the assignment. First a function `NOfit` that calculates the parameters and the residuals for the fit is created.

Question 1.1

The function should have the call `[x_star, r_star] = NOfit(t,y,n)`, but for convenience it is written to also return the design matrix A. The function `NOfit` relies on `get_A` to calculate the actual design matrix. The function `get_A` is used later when the fits are plotted.

```
function [xstar, rstar, A] = NOfit(t, y, n)

    omega = 2*pi/24;
    m = length(t);

    if m ~= length(y)
        error('The length of t and y should match')
    end

    A = get_A(t, n);

    xstar = (A'*A)\(A'*y);
    rstar = y - A*xstar;
end
```

Code Listing 1: `NOfit` function to fit sine, cosines of order `n`

```
function [A] = get_A(t, n)

    if mod(n, 2) == 0
        error('Only odd number of basis functions');
    end

    omega = 2*pi/24;
    m = length(t);
    A = zeros(m, n);
    fns = {@sin, @cos};
```

```
    for i=1:m
        for j=1:n
            % Cycle between sine and cosine
            f = fns{mod(j,2)+1};
            A(i,j) = f(floor(j/2)*omega*t(i));
        end
    end
end
```

Code Listing 2: Helper function `get_A` used to generate design matrix

Question 1.2

See appendix A.1 for code used in this exercise.

The `NOfit` function is now tested. Using the data given in the assignment text a 3rd order cosine is fitted which gives the model.

$$M(\mathbf{x}, t) = 186.81 - 44.94\sin(\omega t) - 93.43\cos(\omega t)$$

To confirm the implementation of `NOfit` the residual 2-norm is checked and it is the expected $||\,\mathbf{r}^*\,||_2 = 292.558$. Since the implementation seems to be right, the fitted model is plotted along with the data in figure 1 .

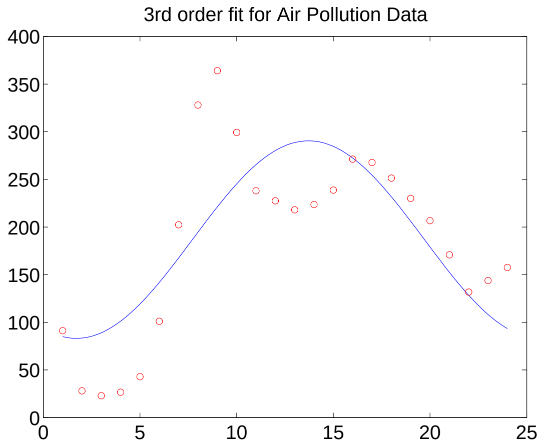

Figure 1: CAPTION!!!

Question 1.3

See appendix A.3 for code used in this exercise.

The optimal order of the fit must be determined. Using the test for random signs and the test for correlation, for the orders $n = 3, 5, 7, 9, 11, 13$ gives the results shown in figure 2.

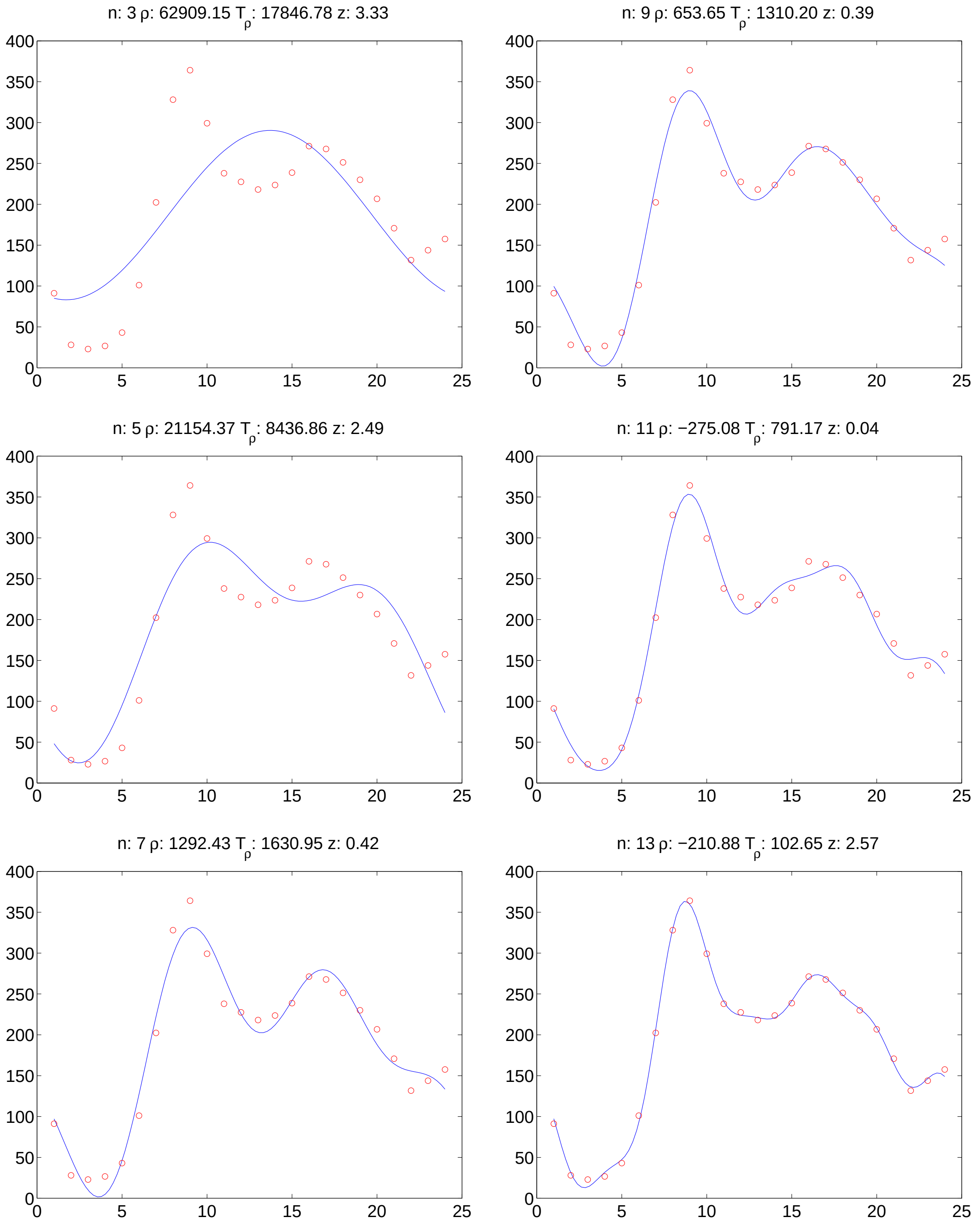

Figure 2: CAPTION!!!

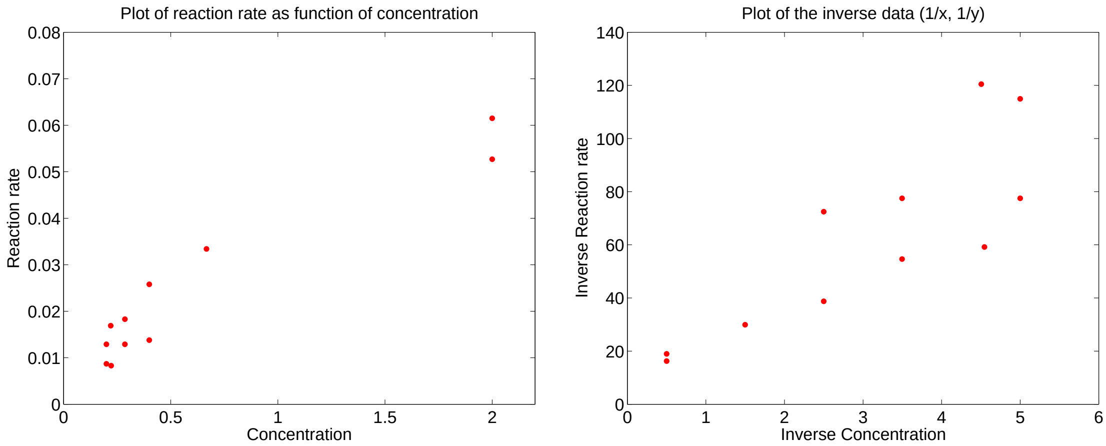

Figure 3: Plot of experimental data (x, y) and inverse data $(1/x, 1/y)$ for question 2

Question 2

In this problem a chemical reaction rate is modelled as a function of the concentration of a substrate. The predicted reaction rate is modelled as

$$\hat{y} = \frac{\theta_1 x}{\theta_2 + x}$$

where x is the concentration and θ_1 and θ_2 are the parameters of interest. Given the 12 measurements of reaction rate and corresponding concentration we then model the measured reaction rate y by

$$y = \hat{y} + e$$

where $e \sim N(0, \sigma^2)$ and σ^2 is unknown.

Question 2.1

See appendix **??** *for code used in this exercise.*

First a plot of the experimental data (x, y) and another plot of the inverse data $(1/x, 1/y)$ are created and shown in figure 3. From the plots it do look as if a straight line could be fitted to the inverse data.

A Appendices

All MATLAB source code is included in the appendices. All the source code including the LaTex code used for the report can also be found at `https://github.com/alphabits/dtu-fall-2011/tree/master/02610/assignment-2`.

A.1 Question 1.2

```
load_ex1;

n = 3;

[xstar, rstar] = NOfit(t, y, n);

file = fopen('../tables/3rd-order-fitted-model.tex', 'w');
fprintf(file, 'M(\\myvec{x}, t) = %.02f %+.02f \\sin(\\omega t) %+.02f \\cos(\\omega t)', xstar);
fclose(file);

plot_fit(t, y, n, xstar, rstar, '3rd order fit for Air Pollution Data');
saveeps('../media/3rd-order-fit.eps');
```

Code Listing 3: ex12.m

A.2 Question 1.3

```
load_ex1;


for n=3:2:13
    [x, r] = NOfit(t, y, n);
    plot_fit_with_res_analysis(t, y, n, x, r);
    saveeps(sprintf('../media/order-determination-%d.eps', n));
end
```

Code Listing 4: ex13.m

A.3 Question 2.1

```
dat = load('../data/reaction-rates.txt');

x = dat(:,1);
y = dat(:,2);

fs = 16;

plot(x, y, 'r.', 'MarkerSize', 16);
```

```
set(gca, 'FontSize', fs);
axis([0 2.2 0 0.08]);
xlabel('Concentration');
ylabel('Reaction rate');
title('Plot of reaction rate as function of concentration');
saveeps('../media/ex21-plot.eps');

plot(1./x, 1./y, 'r.', 'MarkerSize', 16);
set(gca, 'FontSize', fs);
axis([0 6 0 140]);
xlabel('Inverse Concentration');
ylabel('Inverse Reaction rate');
title('Plot of the inverse data (1/x, 1/y)');
saveeps('../media/ex21-plot-inv.eps');

%theta = calc_chemical_reaction_params_linear(x, y);
%yhat = @(x, theta)(theta(1)*x)./(theta(2)+x);
%x_preds = linspace(0,2,100);
%y_preds = yhat(x_preds, theta);

%plot(x_preds, y_preds, 'b-');

% plot(1./x, 1./y, 'ro');
% axis([0 6 0 140]);
```

Code Listing 5: ex21.m

A.4 Helper functions

```
function [] = plot_fit(t, y, n, x, r, plot_title)
    tplot = linspace(1,24,100);
    A = get_A(tplot, n);
    fit = A*x;
    fs = 18;
    set(gca, 'fontsize', fs);
    plot(tplot, fit, '-b', t, y, 'or');
    title(plot_title, 'fontsize', fs);
end
```

Code Listing 6: plot_fit.m

```
function [] = plot_fit_with_res_analysis(t, y, n, x, r)
    z = run_score(r);
    [rho, Trho] = correlation_score(r);
    plot_title = sprintf('n: %d \\rho: %.02f T_\\rho: %.02f z: %.02f', ...
                         n, rho, Trho, z);
    plot_fit(t, y, n, x, r, plot_title);
end
```

Code Listing 7: plot_fit_with_res_analysis.m

References

[1] Jorge Nocedal & Stephen J. Wright, *Numerical Optimization.* Springer Science+Business Media, 2nd Edition, 2006.

[2] Kaj Madsen & Hans Bruun Nielsen, *Introduction to Optimization and Data Fitting.* DTU IMM, 1st Edition, 2010.

[3] Hans Bruun Nielsen, *Checking Gradients.* DTU IMM, 1st Edition, 2000, `http://www2.imm.dtu.dk/~hbn/Software/checkgrad.ps`.