

## Project 5

In the project description we sketch the analysis of the problem we expect you to do. (Obviously you are welcome to do more.) To pass the project a short report should be written and handed in to the project supervisor. In addition the group should present their results in class. The presentation should take about 15 minutes. Please include a short introduction which will facilitate for other students to understand the results of the project. (Do not assume that the audience knows the subject.)

### 1 Introduction - Fatigue of metals

Suppose that a metal structure can endure a static load of  $S$ . However, if the structure breaks down subject to a cyclic load having a peak value noticeably below  $S$ , then the structure is said to suffer from metal fatigue. Particularly, cyclic loads seem to accelerate crack initiation and crack growth. The fatigue crack propagation is a very complicated process. In the initiation phase, most cracks stop to grow while some collapse and create longer cracks that grow with increasing speed. Initiation of cracks usually takes place at the surface at locations where the stress concentration is high.

Among the more spectacular accidents involving fracture due to fatigue are the accidents with the first jet aeroplane in regular traffic (from May 1952), the Comet aeroplane, built by De Havilland. Three of these disintegrated in mid air, and all aeroplanes of this type were withdrawn from traffic. More facts about these crashes are found at the Internet<sup>1</sup>.

In this project you will investigate how a fatigue crack grows under laboratory conditions. You will employ statistical tools called linear regression and bootstrap.

### 2 Fatigue and Fracture Mechanics – Paris' Law

Let consider how an existing crack will grow under constant-amplitude load, i.e. a stress  $s(t) = S \cdot \cos(\omega t)$ . The important quantity is  $N(t) = \frac{\omega}{2\pi} t$  which is the number of stress oscillations in time period  $t$ . Denote by  $A(t)$  the crack length at time  $t$ . It is well known that the crack length  $A(t)$  depends on time only through number of stress oscillations. In fracture mechanics the following relation between  $A$  and  $N$  has been established from experiments:

$$\frac{dA}{dN} = c(A)^k \tag{1}$$

where  $c$  and  $k$  are constants, specific to the material, geometry of specimen, and type of crack. The constants  $c$  and  $k$  are unknown and one wishes to estimate them in laboratory experiments. In 1979, Virkler et al. performed crack-growth measurements on 2024-T3 aluminium-alloy specimens. In total 68 specimens was tested. Each specimen had a 9 mm long initial crack before it was loaded by sinusoidal stress with the same range  $S$ . For each specimen one had measured the number of oscillations  $N$  as a function of crack length  $A$ , giving a matrix

---

<sup>1</sup><http://www.geocities.com/Heartland/Ridge/1747/comet.htm>

$(A_i, N(A_i))$ .<sup>2</sup> Because of this particularity of the data we choose to use an alternative formulation of the Paris law:

$$\frac{dN}{dA} = C(A)^\beta, \quad (2)$$

obviously  $C = 1/c$  while  $\beta = -k$ . Virkler data will be used to validate the model (2).

## 2.1 Virkler Data

Virkler had measured the number of oscillations  $N$  it took for a crack to grow specific length. The data is saved in the file `Virkler`. The first column contain values of  $A$  while the remaining 68 columns contain number of cycles  $N$  it took for a crack in respective specimens to reach the length  $A$  (given in the first column). The easiest way to understand the data is to load it and plot.

```
clear
v = load('virkler.dat');
v=v(1:136,:);
figure(1)
plot(v(:,1),v(:,2:end))
```

We have plotted only the first 136 rows because there was a shift in sampling of the crack length during the experiment. Simply above the threshold cracks grow very fast making it difficult to accurately measure the relation  $N(A)$ . We propose to analyze the model in this region.

In order to check whether the Virkler data confirms the relation (2) let plot  $\frac{dN}{dA}$  as a function of crack length  $A$ . This can be done using the following Matlab script

```
Ai=(v(1:end-1,1)+v(2:end,1))/2;
dA=(Ai(2)-Ai(1));
i=7;
dNdA=diff(v(:,i+1))./dA;
figure(2);
plot(Ai,dNdA)
```

Two observations are immediate; firstly the data seems to support model Eq. (2); secondly the variability (noisiness) of the measurements is getting smaller with crack length. Since it is easier to fit a linear model and it is beneficial to have a similar noise level in the whole range of  $A$ , we propose to take the logarithm of both sides in Eq. (2) then

$$\ln \frac{dN}{dA} = \beta \cdot \ln A + \ln C. \quad (3)$$

This is a linear relation! Let check if the relation holds for the measured data too

```
figure(3)
plot(log(Ai),log(dNdA))
```

Judging from the plot, there seems to be linear relation between  $\ln(A)$  and  $\ln(dN/dA)$ . However, there is also randomness occurring, probably due to variability of material properties along the crack path and measurements errors. We want to model this with tools from probability.

---

<sup>2</sup>More often one is measuring crack length  $A$  as a function of  $N$ .

We postulate that the relation (3) is valid for the expected values, viz.

$$\mathbf{E} \left[ \ln \frac{dA}{dN} \right] = \beta \cdot \ln A + \ln C.$$

Such type of a model is called a linear regression.

### 2.1.1 Linear regression

Denote by  $X_i$  the values of logarithms  $\ln(dN/dA)$  when crack reaches the length  $A_i$ . Here  $A_i$  are known values (measurement errors can be neglected) while  $X_i$  are given by

$$X_i = \beta \cdot \ln A_i + \alpha + \epsilon_i, \quad (4)$$

where  $\epsilon_i$  are iid  $N(0, \sigma^2)$  distributed random variables. The variance  $\sigma^2$  is assumed to be constant but unknown. The problem is to estimate parameters  $\beta$ ,  $\ln C$  and  $\sigma^2$ . We will use the data  $(x_i, \ln A_i)$ ,  $i = 1, \dots, n = 136$ , plotted on Figure 3, to estimate the parameters.<sup>3</sup>

## 2.2 Estimates of Parameters

To estimate the parameters  $\alpha$  and  $\beta$  in the model given in Eq. (4) one can use the Maximum Likelihood method. In Statistics Toolbox in Matlab, there is a routine `regress` which performs least-squares estimation. (For Gaussian errors  $\epsilon_i$  LS and ML methods are equivalent.) Type `help regress`.

We wish to write the relation (4) in a matrix form. Hence let  $\mathbf{b}$  be a column vector of parameters and  $XX$  be the so called design matrix, defining the model then

$$\mathbf{x} = XX * \mathbf{b} + \epsilon.$$

The following Matlab script can be used to find  $\mathbf{b}$ :

```
t=log(Ai); x=log(dNdA);
XX=[ones(length(t),1) t];
b = regress(x,XX);
```

The purpose of the part `ones(length(t),1)` in the call is to handle the so-called intercept (the constant term). The parameter  $\beta$  is given by `b(2)` while  $\alpha$  is found as `b(1)`. From the help text to the routine (`help regress`), you find that confidence intervals for the estimated parameters may be given as output if requested.

Plot the regression line in the same figure as the original data:

```
xest = b(1) + b(2)*t;
% xest = XX*b;
figure(3), hold on
plot(log(Ai),xest,'r.-')
```

## 2.3 Check of Residuals

In the model it is assumed that the errors  $\epsilon_i$ , called also residuals, belonged to a Gaussian distribution; let us therefore examine the residuals. Create a vector `res` of residuals:

```
res = xest - log(dNdA);
```

---

<sup>3</sup> The variable  $t_i = \ln A_i$  is often called the independent variable, or covariate, while  $x_i$  is called the dependent variable.

Plot first the residuals in a probability paper.

```
figure(4)
wnormplot(res)
```

Assuming normality, we can estimate the variance  $\sigma^2$  of  $\epsilon_i$  (and check if zero mean):

```
par7=wnormfit(res)
figure(3)
plot(log(Ai),xest+wnormrnd(par7(1),par7(2),length(xest),1),'g.')
```

(a) Does the Paris law holds for the crack growth in the specimen?

(a) Does the plots contradict the regression model (4)?

We expect answers "yes" and "no" and conclude that the Paris' law models well the results of Virklers experiment at least for the 7th specimen.

Paris law is an empirical equation which can be solved explicitly. Check that with  $A_0 = 9$  and  $\beta \neq -1$

$$N(A) = \frac{1}{\beta + 1} \exp(\alpha) \left( A^{\beta+1} - A_0^{\beta+1} \right).$$

Now the difference equation used to establish the regression gives the following formula for  $N(A_i)$

$$N(A_i) = \exp(\alpha) \sum_{j=1}^i A_j^\beta \Delta A.$$

The deterministic  $N(A)$  function derived from Paris equation can be compared with the observed  $N(A)$ . This is done in the following plot

```
figure(5)
plot(Ai,(Ai.^(b(2)+1)-9^(b(2)+1))*exp(b(1))/(b(2)+1),'r.')
hold on
plot(Ai,cumsum(exp(xest)*dA),'r')
plot(v(:,1),v(:,i))
```

Agreement seems to be very good. Is this good agreement accidental? One way of checking this is to study the variability of a stochastic solution to Paris equation when the residuals  $\epsilon_i$  are included, viz.

$$N(A_i) = \exp(\alpha) \sum_{j=1}^i A_j^\beta \exp(\epsilon_j) \Delta A. \quad (5)$$

Let now simulate three realisation of the random  $N(A)$  processes

```
Nx=length(xest);
plot(Ai,cumsum(exp(xest+wnormrnd(par7(1),par7(2),Nx,1))*dA),'g')
plot(Ai,cumsum(exp(xest+wnormrnd(par7(1),par7(2),Nx,1))*dA),'g')
plot(Ai,cumsum(exp(xest+wnormrnd(par7(1),par7(2),Nx,1))*dA),'g')
```

The plot illustrates for us that the difference between the measured  $N(A)$  and the one derived from the deterministic Paris equation can be quit large. Conclusion.

## 2.4 Checking the regression model for other specimens

Next we turn to investigate whether the model fit data for other specimens. We have shown above that the difference between the measured process  $N(A)$  and the computed from the deterministic Paris law could be large. Hence we wish to test whether the data does not contradict in obvious way the model (5). This is done by simulating 100 realizations of (5) which will describe the variability of  $N(A)$  as predicted by (5). If the observed  $N(A)$  lies in the band we have no reasons to reject the model.

```

par=[];
for i=1:68;
dNdA=diff(v(:,i+1))./dA;
b = regress(log(dNdA),[ones(length(Ai),1) log(Ai)]);
xest = b(1) + b(2)*log(Ai);
res = xest - log(dNdA);
ss=var(res);
par=[par [b' ss]];
figure(6),clf
plot(Ai,cumsum(exp(xest)*dA),'r')
hold on
plot(v(:,1),v(:,i+1),'k')
pause
for j=1:100
    vrnd=cumsum(exp(xest+wnormrnd(0,ss,length(xest),1))*dA);
    plot(Ai,vrnd,'g')
end
plot(Ai,cumsum(exp(xest)*dA),'r.')
plot(v(:,1),v(:,i+1),'.')
axis([5 40 0 3.2*10^5])
pause
end

```

Do the data contradict the model? (Motivate the answer.)

Does the variability of parameters  $\alpha, \beta, \sigma^2$  depend on estimation errors or there are other types of variabilities that affect the results? For example there can be "weaker" and "stronger" components (due to the variability of the material properties along the crack); the range  $S$  can vary between experiments. These suggest that the randomized version of Paris law is needed to describe cracks growth.

Demonstrate that one should model  $\alpha$  and  $\beta$  as random variables, i.e. variability of parameters can not be explained by estimation errors. A very simple test can be done by computing a very wide confidence interval, 99.99% confidence level, for parameters estimated from a specimen chosen at random. If there are specimens having parameters lying outside the interval it contradicts the assumption of deterministic parameters  $\alpha, \beta$ . Please do it using program `regress`.

## 2.5 Reliability problem:

Finally, use parametric bootstrap to estimate the 0.999-quantile of the specimen strength, e.g. the number of oscillations  $N$  needed for cracks to become 36 mm long.

However, first we shall use bootstrap to check if the proposed model seems to reproduce variability in the observed values of  $N$  well.

```
figure(7),clf
hold on
N=v(end,2:end);
Ns=68;
for j=1:100
    Asim_boot=[];
    for i=1:Ns
        n = ceil(68.*rand(1,1));
        xest = par(n,1) + par(n,2)*log(Ai);
        vrnd=cumsum(exp(xest+wnormrnd(0,par(n,3),length(xest),1))*dA);
        Asim_boot=[Asim_boot; vrnd(end)];
    end
    empdistr(Asim_boot,'k')
end
empdistr(N,'r*')
```

What are your conclusions?

Finally use the bootstrap method to estimate the 0.999-quantile of  $N$ .

