Task 1

In the general linear least squares (GLS) method, we write our prediction as a linear combination of *functions* z_i. Similarly to the previous fitting work you have done, we can express this linear combination as a *matrix product*.

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{a} + \boldsymbol{e}$$

$\boldsymbol{Z}$ stores each function z_i evaluated at each x_j, $\boldsymbol{a}$ is the vector of each function's *coefficient*, and $\boldsymbol{e}$ contains the *observed error values*. By defining a total error function S as the sum of squared error, then taking partial derivatives of S, we can derive a similar set of linear equations that can be solved to find the optimal parameters a_i.

$$\boldsymbol{Z}^\top \boldsymbol{Z}\boldsymbol{a} = \boldsymbol{Z}^\top \boldsymbol{y}$$

As this is in the form $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, you can use data to construct $\boldsymbol{Z}^\top\boldsymbol{Z}$ and $\boldsymbol{Z}^\top\boldsymbol{y}$, then use Gauss elimination or Gauss-Jordan elimination to solve for the parameter vector $\boldsymbol{a}$.

i) Fit a model of the form $y_i = a_0 + a_1 e^{-x_i} + a_2 e^{-2x_i}$ to the following data using the above described method. (The expected solution is given in the slides - you may have to click the down arrow on the slides to see the full solution!)

```
x = [-3.0, -2.3, -1.6, -0.9 -0.2, 0.5, 1.2, 1.9, 2.6, 3.3, 4.0]
y = [8.26383742, 6.44045188, 4.74903073, 4.565647, 3.61011683, 3.32743918,
    2.9643915, 1.02239181, 1.09485138, 1.84053372, 1.49110572]
```

Task 2

The matrix $(\boldsymbol{Z}^\top\boldsymbol{Z})^{-1}$ contains statistical information about the coefficients a_i. The *variance* values $\mathrm{Var}(a_i) \propto \left[(\boldsymbol{Z}^\top\boldsymbol{Z})^{-1}\right]_{i,i}$ lie along the diagonal entries and the *covariances*, $\mathrm{Cov}(a_i, a_j) \propto \left[(\boldsymbol{Z}^\top\boldsymbol{Z})^{-1}\right]_{i,j}, i \neq j$, lie in the off-diagonal entries. The variance still needs to be multiplied by the standard error. Letting

$$f(x_i) = a_0 z_0(x_i) + a_1 z_1(x_i) + \cdots + a_{m-1} z_{m-1}(x_i),$$

the standard error squared is given by

$$s^2 = \frac{1}{n-m} \sum_{i=0}^{n-1} (y_i - f(x_i))^2$$

So, $\mathrm{Var}(a_i) = s^2 \left[(\boldsymbol{Z}^\top\boldsymbol{Z})^{\boldsymbol{-1}}\right]_{i,i} = s^2(a_i)$.

i) Fit a function of the form $y_i = a_0 + a_1 x_i + e_i$ to the following data.

```
x = [10.0, 16.3, 23.0, 27.5, 31.0, 35.6, 39.0, 41.5, 42.9, 45.0, 46.0,
45.5, 46.0, 49.0, 50.0]
y = [8.953, 16.405, 22.607, 27.769, 32.065, 35.641, 38.617, 41.095,
43.156, 44.872, 46.301, 47.490, 48.479, 49.303, 49.988]
```

ii) Using the above method, output the variance of each parameter a_i to the terminal.

iii) Let $T = t_{95/2,n-2}$ be the critical value for the t-distribution for 95% confidence with $n-2$ degrees of freedom and let $s(a_i) = \sqrt{s^2(a_i)}$ be the standard deviation of the parameter a_i.

$$\mathbb{P}(a_i \in (a_i - Ts(a_i), a_i + Ts(a_i))) = 0.95$$

Using an appropriate t-distribution, use the parameters' variance values to output 95% confidence intervals to the terminal for each parameter a_i..

* If there is something up with this worksheet, please contact me (Ewan Dalgliesh) at 26192682@students.lincoln.ac.uk or will endeavour to fix any mistakes! (Apologies in advance, of course)