STUDIO 3. THE TANK REACTOR: ARRHENIUS' LAW.

In this exercise we first study how to solve overdetermined systems of equations by *the method of least squares*. Then we continue our study of the tank reactor by using the least squares method to compute the coefficients k_0 and E in the Arrhenius law, by fitting the rate law to measured data.

Important: write a readable report of your work in each studio session. You will need this at the end of the course when you prepare for the written exam; several exam questions will be based directly on the studio work which is only documented in your own notes and the instructions.

Some exercises are called homework in order to save time in the studio classroom.

1. The method of least squares

Consider the linear system of equations

$$Ax = b, \tag{1}$$

where $A \in \mathbf{R}^{m\times n}$, $x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$. If $m > n$ (more equations than unknowns), then the system is "overdetermined" and such a system has no solution in general. Geometrically, the reason for this is that (in general) the vector b lies outside the *range* ("värderummet") of A,

$$R(A) = \{y \in \mathbf{R}^m : y = Ax \text{ for some } x \in \mathbf{R}^n\}.$$

Since the dimension of $R(A)$ is $\le n$ and $m > n$, we realize that the space $R(A)$ does not "fill out" the whole space $\mathbf{R}^m$. Therefore it is likely that a given vector $b \in \mathbf{R}^m$ will lie outside $R(A)$, see Figure 1, and then Ax cannot be equal to b and (1) has no solution.

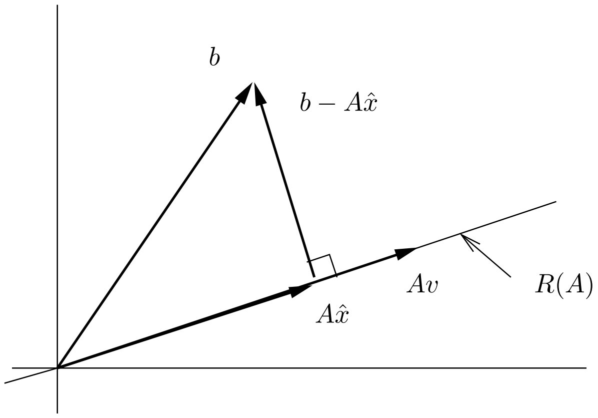

Figure 1. Orthogonal projection onto the range of A.

In this situation we seek an approximate solution which makes the *residual*

$$b - Ax$$

as small as possible. More precisely, we seek a vector $\hat{x} \in \mathbf{R}^n$, which minimizes the square of the norm (length) of the residual:

$$f(\hat{x}) = \min f(x), \quad f(x) = \|b - Ax\|^2. \tag{2}$$

Date: March 11, 2003, Niklas Ericsson, Stig Larsson, Mathematics, Chalmers University of Technology.

Recall the scalar product $\langle x, y\rangle = y^t x$ and the norm $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x^t x}$ of column vectors. We know that there is a unique vector $\hat{y} = A\hat{x} \in R(A)$ such that the distance $\|b - Ax\|$ is minimal, i.e., $\|b - A\hat{x}\| \le \|b - Ax\|$ for all x. The vector $\hat{y} = A\hat{x}$ is the *orthogonal projection* of b onto $R(A)$, see Figure 1. It is characterized by the condition that $b - \hat{y} = b - A\hat{x}$ is orthogonal to all vectors $Av \in R(A)$. This means that

$$0 = \langle b - A\hat{x}, Av\rangle = (Av)^t(b - A\hat{x}) = v^t A^t(b - A\hat{x}) = v^t(A^t b - A^t A\hat{x}).$$

Since this holds for all $v \in \mathbf{R}^n$, we may take $v = A^t b - A^t A\hat{x}$ to get

$$A^t A\hat{x} - A^t b = 0. \tag{3}$$

Thus, we can compute $\hat{x}$ by solving the linear system

$$A^t Ax = A^t b. \tag{4}$$

Note that the coefficient matrix $A^t A$ is $n \times n$ and symmetric. The system (4) has at least one solution (namely $\hat{x}$).

In order to see that the minimization problem (2) is equivalent to solving the linear system (4), we write $x = \hat{x} + v$, $v = x - \hat{x}$, and compute

$$\begin{aligned} f(x) &= f(\hat{x} + v) = \|(b - A\hat{x}) - Av\|^2 = \langle (b - A\hat{x}) - Av, (b - A\hat{x}) - Av\rangle \\ &= \langle b - A\hat{x}, b - A\hat{x}\rangle - 2\langle b - A\hat{x}, Av\rangle + \langle Av, Av\rangle \\ &= \|b - A\hat{x}\|^2 + 2(Av)^t(A\hat{x} - b) + \|Av\|^2 \\ &= f(\hat{x}) + 2v^t A^t(A\hat{x} - b) + \|Av\|^2 \\ &= f(\hat{x}) + 2v^t(A^t A\hat{x} - A^t b) + \|Av\|^2. \end{aligned} \tag{5}$$

Taking (3) into account we get

$$f(x) = f(\hat{x}) + \|Av\|^2 \ge f(\hat{x}).$$

This shows that x minimizes $f(x)$, if and only if $\|Av\| = 0$, in which case $Ax = A\hat{x} + Av = A\hat{x}$ and $A^t Ax = A^t A\hat{x} = A^t b$. Therefore, x minimizes $f(x)$ if and only if x is a solution of (4).

We can also interpret this in terms of the general minimization problem. Recall Taylor's formula:

$$f(x) = f(\hat{x} + v) = f(\hat{x}) + v^t f'(\hat{x}) + \tfrac{1}{2} v^t f''(\hat{x})v + R(x). \tag{6}$$

Noting that $\|Av\|^2 = \langle Av, Av\rangle = (Av)^t(Av) = v^t(A^t A)v$, we re-write (5) as

$$f(x) = f(\hat{x} + v) = f(\hat{x}) + 2v^t(A^t A\hat{x} - A^t b) + v^t(A^t A)v. \tag{7}$$

Comparing (6) with (7), we identify the Jacobi matrix (gradient vector) $f'(\hat{x}) = 2(A^t A\hat{x} - A^t b)$, the Hesse matrix $f''(\hat{x}) = 2A^t A$, and the remainder $R(x) = 0$. Recall that stationary points are given by the system of equations $f'(x) = 2(A^t Ax - A^t b) = 0$, which is the same as (4). Note also that the Hesse matrix is constant (with respect to x) and positive semidefinite: $v^t f''(x)v = 2v^t(A^t A)v = 2\|Av\|^2 \ge 0$.

Exercise 1. Suppose that the variables y and x are related by $y = kx + m$. In order to determine the coefficients k and m we make measurements of y and x:

x	5	6	7	8	9	10
y	19.5888	23.4043	25.5754	29.1231	31.9575	35.8116

This leads to an overdetermined system of the form

$$\begin{gathered} kx_1 + m = y_1 \\ \vdots \\ kx_6 + m = y_6 \end{gathered}$$

or, in matrix form $Av = y$,

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_6 & 1 \end{bmatrix} \begin{bmatrix} k \\ m \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_6 \end{bmatrix}.$$

Solve this system by the least squares method in Matlab. Hint: set up the column vectors `x`, `y` and the matrix `A=[x ones(size(x))]`, then form the matrices `B=A'*A` and `g=A'*y`. Solve the system $Bv = g$ by the command `v=B\g`.

Plot the data points (x_i, y_i) and the fitted function $y = kx+m$ in the same figure. The following commands are useful: `plot(x,y,'or')`, `fplot('ykxm',[x(1) x(6)])`. Here `ykxm.m` is a function file that implements the function $y = kx + m$. Don't forget to declare `global k m` both inside the function file and in the main program.

Actually, Matlab's backslash command `v=A\y` automatically uses the least squares method when the system $Av = y$ is overdetermined. Try this also! □

2. The tank reactor

The rate coefficient depends on the temperature according to the Arrhenius law:

$$k = k_0 \exp(-E/(RT)) \quad \texttt{[s}^{-1}\texttt{]} \tag{8}$$

where R `[8.31 J/(mol K)]` is the gas constant, E `[J/mol]` is the activation energy and k_0 `[s`$^{-1}$`]` is the rate constant of the reaction. The following rates have been measured:

T [K]	343	353	363	373	383	393	403
k [s^{-1}]	$2.8 \cdot 10^{-5}$	$5.6 \cdot 10^{-5}$	$11.2 \cdot 10^{-5}$	$22.4 \cdot 10^{-5}$	$44.8 \cdot 10^{-5}$	$89.6 \cdot 10^{-5}$	$179.2 \cdot 10^{-5}$

The task is now to determine the coefficients k_0 and E by fitting the rate law (8) to these data. Last studio session we wrote (8) in dimensionless form

$$k\tau = \delta e^{\gamma(1-1/X)}, \quad \text{where } \gamma = \frac{E}{RT_f}, \quad \delta = k_0 \tau e^{-\gamma}, \quad X = \frac{T}{T_f}, \quad \tau = \frac{V}{q_{\text{ref}}}. \tag{9}$$

Introducing new variables $r = k\tau$ and $\xi = 1 - 1/X$ we get

$$r = \delta e^{\gamma\xi}. \tag{10}$$

The task is now to fit this function to the given data points (ξ_i, r_i).

Exercise 2. (Linear least squares method.) Form the logarithm of (10) so that you get a linear relation of the form $y = kx + m$, namely,

$$\log(r) = \gamma\xi + \log(\delta). \tag{11}$$

(Note that the natural logarithm is denoted $\log(x)$ in English and in Matlab, but $\ln(x)$ in Swedish.) Solve for γ and δ by using the least squares method as in Exercise 1. Begin by forming column vectors `r`, `X`, `xi` and so on. Plot the data points (X_i, r_i) and the fitted function $r = \delta e^{\gamma(1-1/X)}$ in the same figure. Finally, determine k_0 and E. □

Homework 1. (A nonlinear least squares method.) Alternatively, we can form the residual $r - \delta e^{\gamma\xi}$ from the nonlinear relation (10) and minimize the square of its norm

$$g(\delta, \gamma) = \sum_i \left(r_i - \delta e^{\gamma \xi_i} \right)^2. \tag{12}$$

Write a Matlab function that implements this function and use Matlab's program `fminsearch` to minimize it. Why does this method give a slightly different result? Hint: the Matlab function `norm` may be useful for computing the right-hand side of (12). □

Exercise 3. Insert the new values for δ and γ in your Matlab programs from *Studio 2. The Tank Reactor: Heat Balance.* Repeat all the computations. Let U_1 and U_2 be equal to constant values $\bar{U}_1$ and $\bar{U}_2$. Does the solution $X(s) = \begin{bmatrix} X_1(s) \\ X_2(s) \end{bmatrix}$ approach an equilibrium $\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix}$ as $s \to \infty$? In fact, you should be able to find two equilibrium points by choosing different initial values X_0, say, $X_0 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$ and $X_0 = \begin{bmatrix} 0.5 \\ 1.1 \end{bmatrix}$. □

Next studio session we will look for an equilibrium at $\bar{X}_1 = 0.5$ and analyze the stability of this desired operating point.