

Calculating Kinetic Data by Linear Least-Square Data Fitting Methods for an Elementary Chemical Reaction¹

Becky Hanscam

Advanced Linear Algebra Project

April 25, 2014

1 Introduction

An elementary chemical reaction is where two substrates, A and B , are reacted together to create products, C and D , where



For a chemical reaction to occur, molecules in the gas phase must collide with each other in the correct orientation and with enough energy to break the bonds in the reactants and form the bonds for the product. The minimum amount of energy required for a reaction to occur is the activation energy, E_a . This energy barrier is characteristic for every reaction but the higher the activation energy, the slower the reaction rate [5]. However as the temperature of the conditions in which the reaction is run under increases, the kinetic energy of the molecules of the reactants increases causing more collisions to occur. This increases the ratio of molecules with kinetic energies greater than the activation energy, pushing the reaction rate faster [5]. The rate of a reaction not only depends on the temperature and activation energy, but the frequency of molecule collisions and a rate constant, k . Essentially k is the number of collisions of molecules that will result in a reaction per second at a given temperature. The rate constant is unique to a reaction at a given temperature and does not depend of the concentrations of the reactants; it changes only with temperature and in the presence of a catalyst. Derived from the rate law of a reaction, the Arrhenius equation [1,5,8] relates each of these factors, where T is the temperature of the reaction conditions in Kelvin, R is the gas constant equal to $8.314 \text{ J}/(\text{mol K})$, and A is the frequency factor:

$$k = Ae^{-E_a/RT}. \quad (2)$$

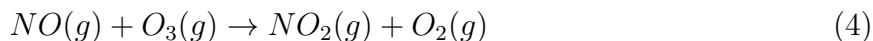
A is the product of the frequency of collisions and a factor expressing the probability that the molecule collides with the correct orientation for a reaction to occur [5]. The Arrhenius equation gives a plot that is nonlinear, but taking the natural logarithm of both sides gives a linearization of the equation [1,5,8]:

$$\ln k = \frac{-E_a}{R} \frac{1}{T} + \ln A \quad (3)$$

in the form $y = mx + b$. A plot of $\ln k$ versus $\frac{1}{T}$ produces a straight lined graph where the slope of the function is $\frac{-E_a}{R}$ and the y -intercept is $\ln A$. This linear relationship allows the calculation of both the activation energy and frequency factor for a reaction.

¹This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>.

Compounds produced in car engines interact with the sunlight in the lower atmosphere forming a mixture of gases that create photochemical smog [5]. These gas molecules react together in a number of different reactions that all follow the general structure of an elementary reaction outlined in (1). One of the reactions in the formation of smog, causing the decomposition of ozone, is the elementary reaction:



where nitrogen dioxide and oxygen gas are produced from nitrogen monoxide and ozone. The kinetics of this reaction and others like it are important to understanding the effects of vehicle exhaust on the atmosphere. From experimental data measuring the rate constant as a function of temperature (Table 1), the plot of k versus T is nonlinear (Figure 1). Using equation (3) a plot of the data with $\ln k$ versus $\frac{1}{T}$ is linear (Figure 2).

Table 1: Temperature Dependence of the Rate Constant in the Formation of Nitrogen Dioxide and Oxygen Gas

T (K)	k ($M^{-1}s^{-1}$)	$\ln k$	$\frac{1}{T}$ (K^{-1})
300	1.21×10^{10}	23.216	3.33×10^{-3}
325	1.67×10^{10}	23.539	3.08×10^{-3}
350	2.20×10^{10}	23.841	2.86×10^{-3}
375	2.79×10^{10}	24.052	2.67×10^{-3}
400	3.45×10^{10}	24.264	2.50×10^{-3}
425	4.15×10^{10}	24.449	2.35×10^{-3}

Note: Data adapted from Gilbert et al.

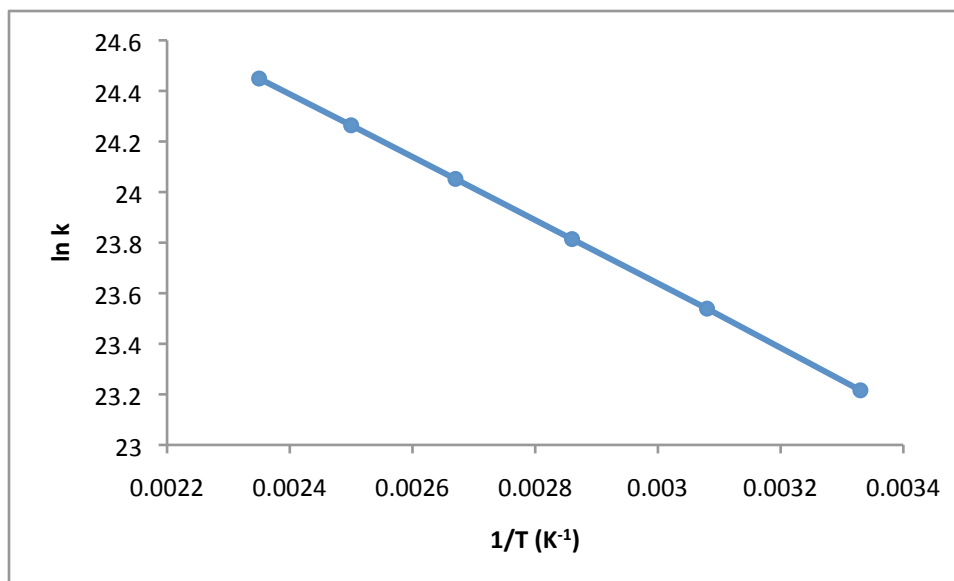


Figure 1: Nonlinear plot of k versus T .

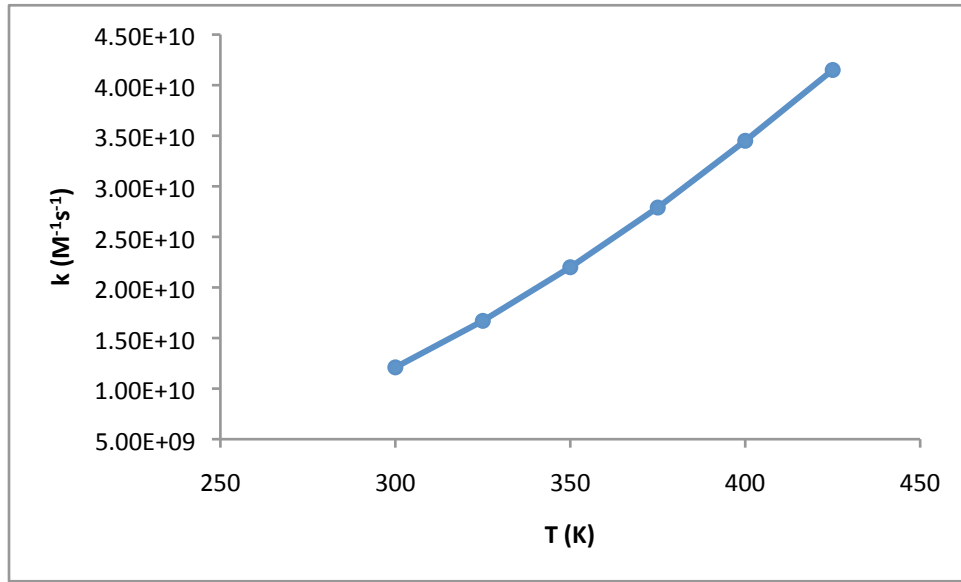


Figure 2: Linear plot of $\ln k$ versus $\frac{1}{T}$.

From the plot of equation (3) the activation energy and frequency factor can be estimated for reaction (4):

$$m = \frac{\Delta y}{\Delta x} = \frac{(23.814 - 24.264)}{(2.86 \times 10^{-3} - 2.50 \times 10^{-3}) K^{-1}} = -1.25 \times 10^3 K$$

$$E_a = -mR = -(-1.25 \times 10^3 K) \times 8.314 (J/mol K) = 1.04 \times 10^4 J/mol = 10.4 kJ/mol$$

$$A = e^{\ln k} = e^{27.41} = 8.0 \times 10^{11}.$$

However the accuracy of the calculations of activation energy and the frequency factor are critical for accurately modeling the behavior of smog and its effects on the environment. Thus mathematical techniques other than estimation are used in these calculations. In an experimental setting, the rate constant can be calculated for a reaction at measured temperatures. Therefore equation (3) can be written as $\mathbf{k} = m\mathbf{T}_0 + b$ where $\mathbf{k} = (\ln k_1, \ln k_2, \dots, \ln k_n)$ and $\mathbf{T}_0 = (\frac{1}{t_1}, \frac{1}{t_2}, \dots, \frac{1}{t_n})$ for n data collections. Then as a matrix-vector product:

$$\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} \Rightarrow \begin{bmatrix} 3.33 \times 10^{-3} & 1 \\ 3.08 \times 10^{-3} & 1 \\ 2.86 \times 10^{-3} & 1 \\ 2.67 \times 10^{-3} & 1 \\ 2.50 \times 10^{-3} & 1 \\ 2.35 \times 10^{-3} & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 23.216 \\ 23.539 \\ 23.841 \\ 24.052 \\ 24.264 \\ 24.449 \end{bmatrix} \Rightarrow T\mathbf{x} = \mathbf{k}.$$

However, because the data points are found experimentally it is unlikely that \mathbf{k} is in the column space of T and so in almost all experiments the system will have no solution. Instead it is possible to compute the least-squares vector that minimizes the square of the distance between the vector and the solution.

Definition [7] A least-squares solution to the system of equations $T\mathbf{x} = \mathbf{k}$ is a vector \mathbf{x} that minimizes $\|\mathbf{k} - T\mathbf{x}\|^2$.

There are multiple computational techniques to find the least-squares solution \mathbf{x} , including QR decomposition via Gram-Schmidt orthogonalization, single value decomposition, Cholesky factorization, and the normal equations.

2 Methods

2.1 Normal Equations

Theorem 2.1. *The least-squares solution to $T\mathbf{x} = \mathbf{k}$ is also a solution to $T^*T\mathbf{x} = T^*\mathbf{k}$, the normal equations, where the function $r(\mathbf{x}) = \|T\mathbf{x} - \mathbf{k}\|^2$ is minimized [7].*

Proof. The vector \mathbf{k} can be decomposed into component vectors where \mathbf{k}_{\parallel} is in the column space of T and \mathbf{k}_{\perp} is orthogonal to \mathbf{k}_{\parallel} and every column of T , $\mathbf{k} = \mathbf{k}_{\parallel} + \mathbf{k}_{\perp}$. Therefore $T^*\mathbf{k}_{\perp} = \mathbf{0}$.

Suppose \mathbf{x} is a least-squares solution. Then

$$\begin{aligned} T^*T\mathbf{x} &= T^*\mathbf{k}_{\parallel} \\ &= T^*(\mathbf{k} - \mathbf{k}_{\perp}) \\ &= T^*\mathbf{k} - T^*\mathbf{k}_{\perp} \\ &= T^*\mathbf{k}. \end{aligned}$$

So \mathbf{x} is a solution to the equation $T^*T\mathbf{x} = T^*\mathbf{k}$. If T is full rank then T^*T and the system of normal equations are nonsingular and \mathbf{x} is a unique least-squares solution [7].

The proof showing that \mathbf{x} minimizes $r(\mathbf{x})$ was presented in class and will not be repeated here. □

In this case, because the rate constant and temperature are the only two variables measured experimentally, T will always have size $m > n$ where n is always 2 and m is equal to the number of data points collected.

Theorem 2.2. *If T is size $m \times n$ with $m \geq n$, then T has full rank if and only if its columns form a linearly independent set.*

Proof. (\Rightarrow) If T has full rank then by definition its rank is equal to n . From basic knowledge we know that n equals the rank plus the nullity of T where nullity is the dimension of the null space of T . In this case, since rank is equal to n , the nullity of T is zero. Therefore T is nonsingular, and from the properties of a nonsingular matrix we know the columns of T form a linearly independent set. (\Leftarrow) If the columns of T are linearly independent then the nullity of T is zero. Then $r(T) + 0 = n$ so the rank of T is equal to the number of columns, n . So by definition T is full rank. □

As shown in the construction of the system $T\mathbf{x} = \mathbf{k}$, the first column of T holds the experimental temperatures and the second column has ones in every entry. Thus there is no scalar that can take the first column to the second, so the columns of T will always form a linearly independent set. Then by Theorem 2.2, T will always have full rank. Because T will always have full rank, the least-squares solution \mathbf{x} will always be unique by Theorem 2.1. For reaction (4) using the normal equations method:

$$T^*T\mathbf{x} = T^*\mathbf{k}$$

$$\begin{bmatrix} 3.33 \times 10^{-3} & 1 \\ 3.08 \times 10^{-3} & 1 \\ 2.86 \times 10^{-3} & 1 \\ 2.67 \times 10^{-3} & 1 \\ 2.50 \times 10^{-3} & 1 \\ 2.35 \times 10^{-3} & 1 \end{bmatrix}^* \begin{bmatrix} 3.33 \times 10^{-3} & 1 \\ 3.08 \times 10^{-3} & 1 \\ 2.86 \times 10^{-3} & 1 \\ 2.67 \times 10^{-3} & 1 \\ 2.50 \times 10^{-3} & 1 \\ 2.35 \times 10^{-3} & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 3.33 \times 10^{-3} & 1 \\ 3.08 \times 10^{-3} & 1 \\ 2.86 \times 10^{-3} & 1 \\ 2.67 \times 10^{-3} & 1 \\ 2.50 \times 10^{-3} & 1 \\ 2.35 \times 10^{-3} & 1 \end{bmatrix}^* \begin{bmatrix} 23.216 \\ 23.539 \\ 23.841 \\ 24.052 \\ 24.264 \\ 24.449 \end{bmatrix}$$

$$\begin{bmatrix} 4.7656 \times 10^{-5} & 0.01679 \\ 0.01679 & 6 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0.40025 \\ 143.334 \end{bmatrix}$$

$$m = -1256.73203263 \Rightarrow E_a = mR = 10.44847012 \frac{kJ}{mol}$$

$$b = 27.405755138 \Rightarrow A = e^b = 7.983038593 \times 10^{11}.$$

2.2 Cholesky Factorization

Similar to the normal equations method is the use of a Cholesky factorization. From the normal equations, T^*T is a symmetric matrix due to the matrix-adjoint product. In the Arrhenius equation the temperatures are measured in Kelvin, where 0 C is 273 K, so the entries of T^*T are always positive.

Definition [6] If $\langle \mathbf{x}, A\mathbf{x} \rangle > 0$ for all \mathbf{x} then A is a symmetric positive definite matrix where $\mathbf{x} \neq 0$.

Therefore, by definition T^*T is a symmetric positive definite matrix.

Theorem 2.3. *If T^*T is symmetric positive definite then there exists a unique upper triangular matrix G with positive diagonal entries such that $T^*T = G^*G$ [6].*

Proof. The first part of this proof was discussed in class but as it is constructive I will repeat it here. Because the matrix T^*T is symmetric positive definite, it can be divided into a block diagonal matrix where the first block is a 1×1 matrix. Using row operations and subsequent column operations T^*T can be converted into the identity matrix. The net effect of the row operations is accumulated in a lower triangular matrix and the net effect of the column operations is stored in an upper triangular matrix.

$$\begin{aligned}
T^*T = A &= \left[\begin{array}{c|c} a & \mathbf{y}^* \\ \mathbf{y} & B \end{array} \right] \\
&= \left[\begin{array}{c|c} \sqrt{a} & \mathbf{0}^* \\ \frac{1}{\sqrt{a}}\mathbf{y} & I \end{array} \right] \left[\begin{array}{c|c} 1 & \mathbf{0}^* \\ \mathbf{0} & B - \frac{1}{a}\mathbf{y}\mathbf{y}^* \end{array} \right] \left[\begin{array}{c|c} \sqrt{a} & \frac{1}{\sqrt{a}}\mathbf{y}^* \\ \mathbf{0} & I \end{array} \right] \\
&= G_1^* A_1 G_1
\end{aligned}$$

Since T^*T is a Hermitian matrix, it will always have real diagonal entries. So a will always be positive. Repeated decompositions on the matrix $B - \frac{1}{a}\mathbf{y}\mathbf{y}^*$ will eventually give the identity matrix. The entry in the upper left corner of this matrix will also always be positive because T^*T is symmetric positive definite:

$$a = \langle \mathbf{e}_2, A_1 G_1^{-1} \mathbf{e}_2 \rangle > 0 \quad \text{where } \mathbf{x} = G_1^{-1} \mathbf{e}_2$$

After n iterations:

$$A = G_n^* \dots G_2^* G_1^* I G_1 G_2 \dots G_n = G^* G$$

Then [2]

$$T^*T = G^*G \rightarrow G^*G\mathbf{x} = T^*\mathbf{k}$$

□

For reaction (4) using the Cholesky factorization, we can see

$$\begin{aligned}
T^*T &= \begin{bmatrix} 4.7656 \times 10^{-5} & 0.01679 \\ 0.01679 & 6 \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{4.7656 \times 10^{-5}} & 0 \\ \frac{0.01679}{\sqrt{4.7656 \times 10^{-5}}} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 6 - \frac{0.01679^2}{4.7656 \times 10^{-5}} \end{bmatrix} \begin{bmatrix} \sqrt{4.7656 \times 10^{-5}} & \frac{0.01679}{\sqrt{4.7656 \times 10^{-5}}} \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{4.7656 \times 10^{-5}} & 0 \\ \frac{0.01679}{\sqrt{4.7656 \times 10^{-5}}} & \sqrt{6 - \frac{0.01679^2}{4.7656 \times 10^{-5}}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{4.7656 \times 10^{-5}} & \frac{0.01679}{\sqrt{4.7656 \times 10^{-5}}} \\ 0 & \sqrt{6 - \frac{0.01679^2}{4.7656 \times 10^{-5}}} \end{bmatrix} \\
&= \begin{bmatrix} 6.90 \times 10^{-3} & 0 \\ 2.4322 & 0.29093 \end{bmatrix} \begin{bmatrix} 6.90 \times 10^{-3} & 2.4322 \\ 0 & 0.29093 \end{bmatrix} \\
&= G^*G.
\end{aligned}$$

We solve the system $G^*\mathbf{w} = T^*\mathbf{k}$ for \mathbf{w} by augmenting the matrix G^* with the vector \mathbf{k} and row reducing. Then we can find the least-squares solution by solving the system $G\mathbf{x} = \mathbf{w}$ for \mathbf{x} by augmenting the matrix G with the vector w and row reducing. This produces

$$\begin{aligned} m &= -1256.74352341 \Rightarrow E_a = mR = 10.44856565 \frac{kJ}{mol} \\ b &= 27.4057876928 \Rightarrow A = e^b = 7.983289484 \times 10^{11}. \end{aligned}$$

2.3 QR Decomposition

The method of QR decomposition via the Gram-Schmidt procedure gives the least-squares solution to any type of coefficient matrix for a system of equations, making it a particularly useful method.

Theorem 2.4. (Gram-Schmidt procedure) [2] Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_p\}$ is a linearly independent set of vectors. Define the vectors \mathbf{u}_i , $1 \leq i \leq p$ by

$$\mathbf{u}_i = \mathbf{v}_i - \frac{\mathbf{u}_1^* \mathbf{v}_i}{\mathbf{u}_1^* \mathbf{u}_1} \mathbf{u}_1 - \frac{\mathbf{u}_2^* \mathbf{v}_i}{\mathbf{u}_2^* \mathbf{u}_2} \mathbf{u}_2 - \frac{\mathbf{u}_3^* \mathbf{v}_i}{\mathbf{u}_3^* \mathbf{u}_3} \mathbf{u}_3 - \dots - \frac{\mathbf{u}_{i-1}^* \mathbf{v}_i}{\mathbf{u}_{i-1}^* \mathbf{u}_{i-1}} \mathbf{u}_{i-1}$$

Then $T = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors, and $\langle T \rangle = \langle S \rangle$.

Theorem 2.5. [2, 6] Suppose that T is an $m \times n$ matrix of rank n . Then there exists an $m \times n$ matrix Q whose columns form an orthonormal set, and an upper-triangular matrix R of size n with positive diagonal entries, such that $T = QR$.

Proof. The Gram-Schmidt procedure takes any matrix and produces a set of orthogonal vectors from the columns of the matrix. Each vector of the resulting set is a linear combination of the columns of the original matrix. The coefficients of each linear combination are stored in the upper triangular matrix R . Then when each orthogonal vector is scaled by the inverse of its norm to become orthonormal, the matrix R stores the operations:

$$\begin{aligned} T &= [\mathbf{t}_1 | \mathbf{t}_2] \\ &= [\mathbf{u}_1 | \mathbf{u}_2] \begin{bmatrix} 1 & \frac{-\mathbf{t}_1^* \mathbf{t}_2}{\mathbf{t}_1^* \mathbf{t}_1} \\ 0 & 1 \end{bmatrix}^{-1} && \text{Gram-Schmidt on } \mathbf{t}_1 \text{ and } \mathbf{t}_2 \\ &= [\mathbf{q}_1 | \mathbf{q}_2] \begin{bmatrix} \frac{1}{\|\mathbf{u}_1\|} & \frac{-\mathbf{t}_1^* \mathbf{t}_2}{\mathbf{t}_1^* \mathbf{t}_1} \\ 0 & \frac{1}{\|\mathbf{u}_2\|} \end{bmatrix}^{-1} && \mathbf{u}_1 \text{ and } \mathbf{u}_2 \text{ scaled by their norm} \\ &= QR, \end{aligned}$$

where Q has orthonormal columns and R is square upper triangular with full rank. From the normal equations, we can see

$$\begin{aligned} T^* T \mathbf{x} &= T^* \mathbf{k} \\ R^* Q^* Q R \mathbf{x} &= R^* Q^* \mathbf{k} \\ R \mathbf{x} &= Q^* \mathbf{k}. \end{aligned}$$

Then the system $R\mathbf{x} = Q^*\mathbf{k}$ is nonsingular when T has full rank [2]. \square

This method will always give a unique least-squares solution because, by Theorem 2.2, T will always have full rank. For reaction (4), using the QR decomposition,

$$\begin{bmatrix} 3.33 \times 10^{-3} & 1 \\ 3.08 \times 10^{-3} & 1 \\ 2.86 \times 10^{-3} & 1 \\ 2.67 \times 10^{-3} & 1 \\ 2.50 \times 10^{-3} & 1 \\ 2.35 \times 10^{-3} & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{-\mathbf{t}_1^*\mathbf{t}_2}{\mathbf{t}_1^*\mathbf{t}_1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3.33 \times 10^{-3} & -0.1732 \\ 3.08 \times 10^{-3} & -0.0851 \\ 2.86 \times 10^{-3} & -0.0076 \\ 2.67 \times 10^{-3} & 0.0593 \\ 2.50 \times 10^{-3} & 0.1192 \\ 2.35 \times 10^{-3} & 0.1721 \end{bmatrix} \quad TR_2 = U$$

$$\begin{bmatrix} 3.33 \times 10^{-3} & 1 \\ 3.08 \times 10^{-3} & 1 \\ 2.86 \times 10^{-3} & 1 \\ 2.67 \times 10^{-3} & 1 \\ 2.50 \times 10^{-3} & 1 \\ 2.35 \times 10^{-3} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\|\mathbf{u}_1\|} & \frac{-\mathbf{t}_1^*\mathbf{t}_2}{\mathbf{t}_1^*\mathbf{t}_1} \\ 0 & \frac{1}{\|\mathbf{u}_2\|} \end{bmatrix} = \begin{bmatrix} 0.4824 & -0.5954 \\ 0.4462 & -0.2926 \\ 0.4143 & -0.0262 \\ 0.3868 & 0.2039 \\ 0.3621 & 0.4098 \\ 0.3404 & 0.5914 \end{bmatrix} \quad TR_1 = Q$$

$$T = QR_1^{-1} = \begin{bmatrix} 0.4824 & -0.5954 \\ 0.4462 & -0.2926 \\ 0.4143 & -0.0262 \\ 0.3868 & 0.2039 \\ 0.3621 & 0.4098 \\ 0.3404 & 0.5914 \end{bmatrix} \begin{bmatrix} 6.9034 \times 10^{-3} & 2.4322 \\ 0 & 0.2909 \end{bmatrix} = QR.$$

We compute the matrix vector product $Q^*\mathbf{k} = \mathbf{b}$ and then solve the system $R\mathbf{x} = \mathbf{b}$ for \mathbf{x} by augmenting the matrix R with the vector \mathbf{b} and row reducing, producing

$$m = -1256.73203263 \Rightarrow E_a = mR = 10.44847012 \frac{kJ}{mol}$$

$$b = 27.405755138 \Rightarrow A = e^b = 7.983038593 \times 10^{11}.$$

The QR decomposition via the Gram-Schmidt orthogonalization procedure is successful with any matrix as the input. So the successful calculation of a solution is not dependent on the entries of the matrix. Therefore it is stable in the presence of rounding errors and a valuable computational tool for experimental results in general.

2.4 Singular Value Decomposition

Theorem 2.6. *If T is a real $m \times n$ matrix then there exists orthogonal matrices*

$$U = [\mathbf{u}_1 | \dots | \mathbf{u}_m] \text{ and } V = [\mathbf{v}_1 | \dots | \mathbf{v}_n]$$

, where U is size m and V is size n , such that $T = USV^*$. S is a diagonal matrix with diagonal entries $\sqrt{\delta_1}, \dots, \sqrt{\delta_n}$, where $\delta_1, \dots, \delta_n$ are eigenvalues of the matrix T^*T [6].

Proof. This proof was discussed in class and so will not be repeated here. \square

For reaction (4) using the singular value decomposition, we see the eigenvalues of T^*T , δ_1, δ_2 , are $\{6.72278 \times 10^{-7}, 6\}$. Thus, the singular values of T are $s_1 = \sqrt{\delta_1} = 8.199 \times 10^{-4}$ and $s_2 = \sqrt{\delta_2} = 2.4495$. Placing the singular values as the diagonal entries of a matrix, gives

$$S = [s_1 \mathbf{e}_1 | s_2 \mathbf{e}_2] = \begin{bmatrix} 8.199 \times 10^{-4} & 0 \\ 0 & 2.4495 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The eigenvectors for δ_1 and δ_2 of T^*T , \mathbf{x}_1 , and \mathbf{x}_2 , expressed as columns of a matrix, are

$$V^* = [\mathbf{x}_1 | \mathbf{x}_2]^* = \begin{bmatrix} -0.999996 & 0.002798 \\ -0.002798 & -0.999996 \end{bmatrix}.$$

The first two columns of the unitary matrix are $\mathbf{y}_1 = \frac{1}{\sqrt{\delta_1}} T \mathbf{x}_1$ and $\mathbf{y}_2 = \frac{1}{\sqrt{\delta_2}} T \mathbf{x}_2$ and then the eigenvectors of TT^* for the zero eigenvalue, \mathbf{y}_3 , \mathbf{y}_4 , \mathbf{y}_5 , and \mathbf{y}_6 , are the remaining columns, so that

$$U = [\mathbf{y}_1 | \mathbf{y}_2 | \mathbf{y}_3 | \mathbf{y}_4 | \mathbf{y}_5 | \mathbf{y}_6] \\ = \begin{bmatrix} -0.6484209 & -0.4082489 & -0.6425508 & -0.3461808 & -0.0027296 & -0.0339307 \\ -0.3435253 & -0.4082489 & 0.6060519 & 0.3245593 & -0.4845332 & -0.2458926 \\ -0.0752094 & -0.4082484 & 0.3352808 & 0.1676051 & 0.5944281 & 0.7265551 \\ 0.1565180 & -0.4082481 & 0.1014331 & 0.3999138 & -0.0264467 & -0.5969384 \\ 0.3638531 & -0.4082479 & -0.1077991 & -0.7408139 & 0.4112590 & 0.2215388 \\ 0.5467958 & -0.4082478 & -0.2924158 & 0.1949166 & -0.4919776 & -0.0713323 \end{bmatrix}.$$

We solve the system $SV^* \mathbf{x} = U^* \mathbf{k}$ for \mathbf{x} by augmenting the matrix SV^* with the vector $U^* \mathbf{k}$ and row reducing. However, the vector $U^* \mathbf{k}$ is not in the column space of SV^* so the system is inconsistent, giving a 6×3 matrix with ones on the diagonal and numbers very close to zero everywhere else. In order to find the least-squares solution we write the system in terms of the normal equations. Let $C = SV^*$ and $\mathbf{b} = U^* \mathbf{k}$, then solve the system $C^* C \mathbf{x} = C^* \mathbf{b}$ for the least-squares solution \mathbf{x} . This gives

$$m = -1256.73203461 \Rightarrow E_a = mR = 10.44847014 \frac{kJ}{mol} \\ b = 27.4057551435 \Rightarrow A = e^b = 7.983038637 \times 10^{11}.$$

The matrix T is full rank and so the diagonal system is nonsingular and requires the use of other least-squares methods. In other applications where T is rank deficient, its singular value decomposition produces a system of equations that gives the least-square solution directly, without the use of a second least-square method.

3 Conclusion

Table 2: Results of Various Least-squares Methods for the Calculation of the Activation Energy and Frequency Factor of Reaction (4)

Calculation Method	E_a (kJ/mol)	A
Estimation	10.4	8.0×10^{11}
Normal Equations	10.44847012	$7.983038593 \times 10^{11}$
QR Decomposition	10.44847012	$7.983038593 \times 10^{11}$
Singular Value Decomposition	10.44847014	$7.983038637 \times 10^{11}$
Cholesky Factorization	10.44856565	$7.983289484 \times 10^{11}$

Each method of least-squares calculation has its own benefits and drawbacks. Some are more advantageous or accurate in certain situations than others. However as shown in Table 2, each method calculates the activation energy and frequency factor much more accurately than a direct estimation ever could. The solutions to all the least-squares methods are equal through the third decimal place for both the activation energy and the frequency factor. The normal equations and the QR decomposition methods yield, surprisingly, equal results. This is because the QR decomposition of T via the Gram-Schmidt orthogonalization procedure preserves the values of its entries exactly when calculated over the field RDF. Therefore the QR decomposition method is essentially the same as the normal equations method. One major difference however, is that the QR decomposition is much more stable when there are rounding errors in the data. When errors in rounding are not an issue, the normal equations method is preferred because there are significantly less computations involved than in the QR decomposition method. So when the speed of the computation is the only major consideration, the normal equations and Cholesky factorization methods are favored as they compute solutions much more rapidly, especially with large data sets. The singular value decomposition method is also heavy computationally but is the only method that holds in cases where T is rank deficient. These methods are extremely valuable both in the information they provide and the rapidness and flexibility of the calculation.

In agreement from all the least-squares calculations, the activation energy of reaction (4) is 10.448 kJ/mol and the frequency factor is 7.983×10^{11} . This is a very low energy barrier as seen in Figure 3, low enough that the reaction occurs spontaneously. This helps explain the depletion of ozone in the upper atmosphere as it reacts easily, if not spontaneously, with the nitrogen monoxide emitted from car engines. It is very important, especially in the modern world, to have the ability to accurately monitor the behavior of photochemical smog and chemical interactions with ozone in the atmosphere.

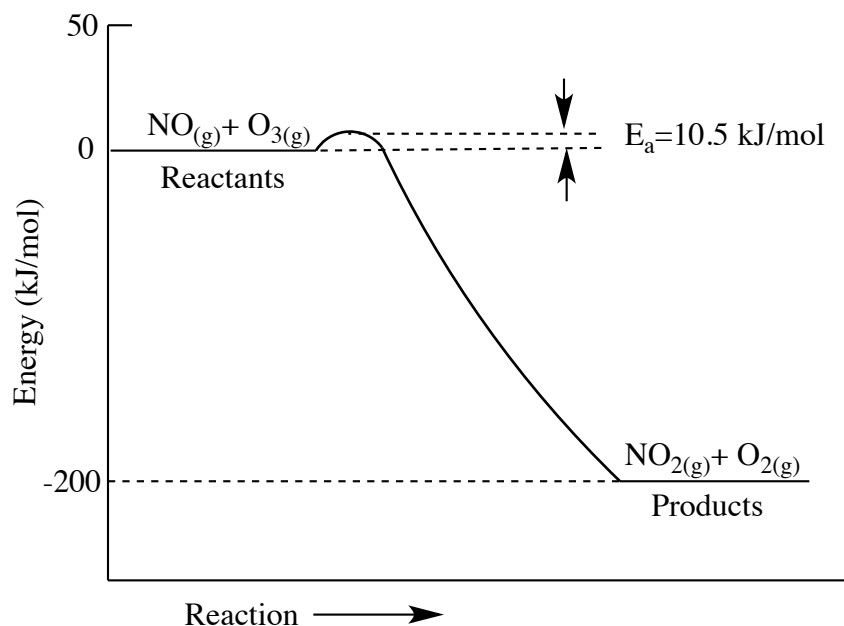


Figure 3: Energy profile for reaction (4).

4 References

- [1] Alanazi, A.; Brown, T.; Kendell, S.; Miron, D.; Munshi, A. *Time-Varying Flexible Least Squares for Thermal Desorption of Gases*. International Journal of Chemical Kinetics: Vol 45, Issue 6; 2013; pp 374-386.
- [2] Bau, D. III; Trefethen, L. *Numerical Linear Algebra*; Siam: 1997; pp 77-85.
- [3] DaCosta, H.; Fan, M. *Rate Constant Calculation for Thermal Reactions: Methods and Applications*; Wiley: Hoboken; 2012.
- [4] Garrido, M.; Larrechi, M.; Rius, F. *Multivariate curve resolution-alternating least squares and kinetic modeling applied to near-infrared data from curing reactions of epoxy resins: mechanistic approach and estimation of kinetic rate constants*. Applied Spectroscopy; Vol 60, Issue 2; 2006; pp 174-181.
- [5] Gilbert, T.; Kirss, R.; Foster, N.; Davies, G. *Chemistry: The Science in Context*; W.W. Norton and Company: New York, 2004; pp 702-761.
- [6] Golub, G.; Van Loan, C. *Matrix Computations*, 4th edition; Johns Hopkins University:

Baltimore, 2013; pp 76-80, 163-4, 246-250, 262-4.

[7] Sadun, L.; *Applied Linear Algebra: The Decoupling Principle*, 2nd Edition; American Mathematical Society: 2008; pp 143, 167-181.

[8] Sundberg, R.; *Statistical aspects on fitting the Arrhenius equation*; Chemometrics and Intelligent Laboratory Systems; Vol 41; 1998; pp 249-252.