Series, Complex Variables,
and
Linear Algebra and Least Squares

ChE 132A handout

This handout collects three short background topics used frequently in chemical engineering courses: infinite series, basic complex variables, and linear algebra with least squares.

1 Series, convergence, and the ratio test

Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers. The associated *infinite series* is written as

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots$$

To decide whether this infinite sum makes sense, we look at its *partial sums*

$$s_N = \sum_{n=1}^{N} a_n \tag{1}$$

We say that the series $\sum_{n=1}^{\infty} a_n$ *converges* to s if the sequence of partial sums $(s_N)_{N=1}^{\infty}$ converges to s

$$\lim_{N\to\infty} s_N = s \tag{2}$$

In that case we write

$$\sum_{n=1}^{\infty} a_n = s$$

If the partial sums do not approach a finite limit, then the series *diverges*.

Remark. *A necessary condition for convergence is that $a_n \to 0$ as $n \to \infty$. To see this, assume that a_n does not converge to zero, which means $s_{n+1} - s_n = a_n$ does not converge to zero, which means the sequence (s_n) does not converge. The condition $a_n \to 0$ is necessary, but not sufficient for series convergence.*

Example 1.1 (A convergent geometric series)**.** Consider

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$$

The partial sums are

$$s_N = 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \left(\frac{1}{2}\right)^N = \frac{1-(1/2)^{N+1}}{1-1/2} = 2 - \left(\frac{1}{2}\right)^N$$

Since $\left(\frac{1}{2}\right)^N \to 0$, we obtain $s_N \to 2$. Therefore,

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 2$$

Example 1.2 (A divergent geometric series)**.** Consider

$$\sum_{n=0}^{\infty} 2^n = 1 + 2 + 4 + 8 + \cdots$$

Here the partial sums are

$$s_N = 1 + 2 + 4 + \cdots + 2^N = 2^{N+1} - 1$$

Since $2^{N+1} \to \infty$, the series diverges.

Example 1.3 (A series that diverges because its terms do not go to zero)**.** Consider

$$\sum_{n=1}^{\infty} \frac{n}{n+1}$$

Its terms satisfy

$$\frac{n}{n+1} \to 1 \neq 0$$

Therefore the series diverges.

The ratio test

Theorem 1.1 (Ratio test)**.** *Let* $\sum_{n=1}^{\infty} a_n$ *be a series with* $a_n \neq 0$ *for all sufficiently large n, and suppose the limit*

$$R = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \tag{3}$$

exists. Then:

- *(a) If* $R < 1$*, the series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely.*
- *(b) If* $R > 1$*, or if* $R = \infty$*, the series diverges.*
- *(c) If* $R = 1$*, the ratio test is inconclusive.*

Example 1.4 (A convergent series by the ratio test)**.** Consider

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

Here $a_n = 1/n!$, so

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{1/(n+1)!}{1/n!} = \frac{1}{n+1}$$

Thus

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = 0 < 1$$

so the series converges.

Example 1.5 (A divergent series by the ratio test)**.** Consider

$$\sum_{n=1}^{\infty} \frac{n!}{3^n}$$

Here $a_n = n!/3^n$, so

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)!/3^{n+1}}{n!/3^n} = \frac{n+1}{3}$$

Hence

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \infty > 1$$

so the series diverges.

Two important cases where the ratio test is inconclusive

Example 1.6 (The harmonic series)**.** Consider

$$s = \sum_{n=1}^{\infty}\frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots$$

If $a_n = 1/n$, then

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1/(n+1)}{1/n} = \frac{n}{n+1}$$

Therefore,

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$$

So the ratio test is inconclusive. Nevertheless, this series is known to diverge.

Example 1.7 (The series of reciprocal squares)**.** Consider

$$s = \sum_{n=1}^{\infty}\frac{1}{n^2} = 1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{3}\right)^2 + \cdots$$

If $a_n = 1/n^2$, then

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1/(n+1)^2}{1/n^2} = \left(\frac{n}{n+1}\right)^2$$

Therefore,

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$$

Again the ratio test is inconclusive. However, this series does converge, and in fact

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6} \tag{4}$$

2 Complex variables

The basic imaginary unit is denoted by i and satisfies

$$i^2 = -1$$

A *complex number* is a number of the form

$$z = a + ib$$

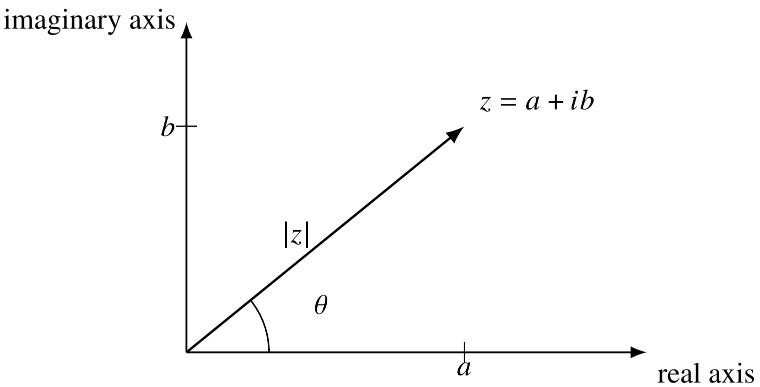

Figure 1: Representation of a complex number in the complex plane.

where a and b are real numbers. The real and imaginary parts of z are

$$\mathrm{Re}(z) = a \qquad \mathrm{Im}(z) = b$$

Two complex numbers $z_1 = a_1 + ib_1$, $z_2 = a_2 + ib_2$ are equal if and only if their real parts are equal and their imaginary parts are equal

$$z_1 = z_2 \quad \text{if and only if} \quad a_1 = a_2 \text{ and } b_1 = b_2$$

The *complex conjugate* of $z = a + ib$ is obtained by changing the sign of the imaginary part

$$\overline{z} = a - ib$$

A complex number can be viewed as a point (a, b) in the complex plane, as shown in Figure 1.

Polar form

The *modulus* of $z = a + ib$ is

$$|z| = \sqrt{z\,\overline{z}} = \sqrt{a^2 + b^2}$$

The *argument* of z is the angle θ that the vector from the origin to z makes with the positive real axis. Since $\tan\theta = b/a$ as shown in Figure 1, a convenient formula is

$$\theta = \arg z = \tan^{-1}(b/a) \qquad \tan\theta = b/a$$

provided the correct quadrant is chosen. In polar coordinates, a complex number can then be written as

$$z = r(\cos\theta + i\sin\theta) \tag{5}$$

with $r = |z|$. Using Euler's formula (see (7)), this can also be written as

$$z = re^{i\theta} \tag{6}$$

Algebra of complex numbers

Addition and subtraction are performed by combining real parts and imaginary parts separately. Let $z_1 = a_1 + ib_1, z_2 = a_2 + ib_2$. Then

$$z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)$$

Multiplication follows the same algebraic rules as for polynomials, together with $i^2 = -1$

$$z_1 z_2 = a_1 a_2 + ia_1 b_2 + ib_1 a_2 + i^2 b_1 b_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2)$$

To divide by z_2, multiply numerator and denominator by the conjugate $\overline{z}_2$

$$\frac{z_1}{z_2} = \frac{z_1 \overline{z}_2}{z_2 \overline{z}_2} = \frac{(a_1 + ib_1)(a_2 - ib_2)}{(a_2 + ib_2)(a_2 - ib_2)} = \frac{(a_1 a_2 + b_1 b_2) + i(-a_1 b_2 + b_1 a_2)}{a_2^2 + b_2^2}$$

Therefore,

$$\frac{z_1}{z_2} = \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + i\frac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2}$$

Using the polar form (6), multiplication and division become much simpler

$$z_1 z_2 = r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$
$$\frac{z_1}{z_2} = \frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)}$$

Also in polar form, repeated multiplication gives

$$z^n = r^n e^{in\theta} = r^n(\cos n\theta + i \sin n\theta)$$

which is known as *De Moivre's theorem.*

Euler's formula

A central identity in complex variables is

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{7}$$

which is illustrated in Figure 2. Substituting (7) into (5) gives the exponential polar form (6).

Example 2.1 (Deriving Euler's formula from the Taylor series)**.** Start with the Taylor series of the exponential function

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Note that this series converges for all x. Now replace x by $i\theta$

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \cdots$$

Using the pattern

$$i^0 = 1 \qquad i^1 = i \qquad i^2 = -1 \qquad i^3 = -i \qquad i^4 = 1$$

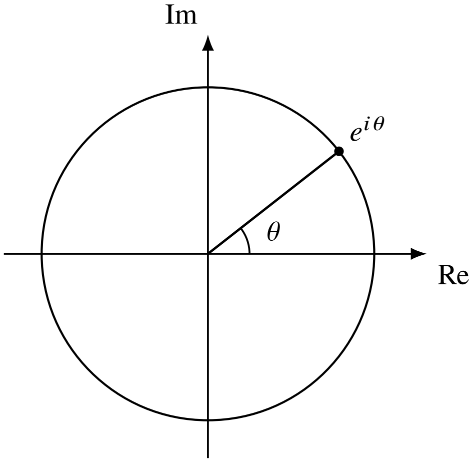

Figure 2: A geometric interpretation of Euler's formula on the unit circle

we obtain

$$e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \cdots$$

Group the real terms and imaginary terms

$$e^{i\theta} = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right)$$

The two series in parentheses are exactly the Taylor series for cosine and sine

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots$$

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots$$

Note that these series converge for all θ. Therefore,

$$e^{i\theta} = \cos\theta + i\sin\theta$$

This establishes Euler's formula.

Euler's formula immediately gives the useful identities

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Defining the hyperbolic functions in the standard way for real-valued x

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad \sinh x = \frac{e^x - e^{-x}}{2}$$

and extending these to complex argument enables us to the see the close connection between the trigonmetric and hyperbolic functions

$$\cosh iz = \cos z \qquad \sinh iz = i\sin z$$

We also see that both trigonometric and hyperbolic functions are simply linear combinations of the exponential functions using different signs in the argument. These "repackagings" of the exponential will prove especially useful when solving linear second-order differential equations in the upcoming lectures.

3 Linear algebra and least squares

Consider the three linear equations in the three unknowns x_1, x_2, and x_3

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

It is convenient to write these equations in matrix form. Define

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Then the system becomes

$$Ax = b \tag{8}$$

Equation (8) is simply a compact way to write the three scalar equations at once.

Norm and inner product

For vectors like x and b, we need a measure of their size and a means to calculate the angle between them. For size, we use the following norm (Euclidean norm, 2-norm) defined by

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

The inner product (dot product) between two vectors x and y is defined by

$$(x, y) = \sum_{i=1}^{n} x_i y_i$$

So we see that this inner product is related to the norm by $(x, x) = \|x\|^2$.

The angle between two (nonzero) vectors is then related to the inner product by

$$(x, y) = \|x\| \, \|y\| \cos\theta \qquad \cos\theta = \frac{(x, y)}{\|x\| \, \|y\|}$$

where θ is the angle between vectors x and y as shown in Figure 3. So if the two vectors x and y are orthogonal, $\cos\theta = 0$ we have that

$$(x, y) = 0 \qquad x \text{ and } y \text{ are orthogonal}$$

We will make extensive use of orthogonality of *functions* as well as vectors in the upcoming lectures.

Linear independence and rank

A collection of vectors $v_1, \dots, v_k$ is called *linearly independent* if the only solution of

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0$$

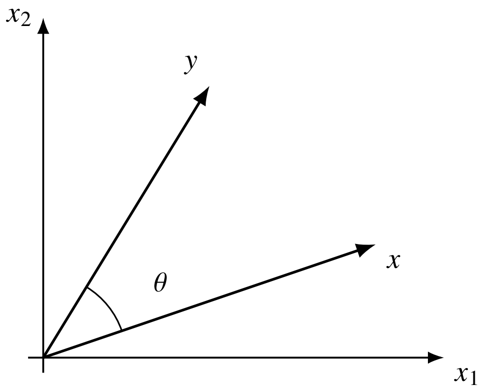

Figure 3: Two vectors x and y and the angle θ between them.

is

$$c_1 = c_2 = \cdots = c_k = 0$$

If there is a nontrivial choice of coefficients that gives zero, then the vectors are *linearly dependent*.

For a matrix A, we can ask whether its columns are linearly independent and whether its rows are linearly independent.

The *rank* of a matrix A, written $\operatorname{rank}(A)$, is the number of linearly independent columns of A. Equivalently, it is the number of linearly independent rows of A. These two numbers are always the same.

For a 3×3 matrix:

- the rows are linearly independent if and only if $\operatorname{rank}(A) = 3$,
- the columns are linearly independent if and only if $\operatorname{rank}(A) = 3$.

So for a square 3×3 matrix, row independence and column independence are equivalent.

When does $Ax = b$ have a solution, and when is it unique?

For a general matrix equation (8), two questions are important:

(a) for a given right-hand side b, does a solution exist,

(b) if a solution exists, is it unique?

The key ideas are:

- $Ax = b$ can be solved for *every* vector b if and only if the rows of A are linearly independent,
- the solution, when it exists, is unique if and only if the columns of A are linearly independent.

For a square 3×3 system these two conditions are equivalent. In that case, if A has rank 3, then A is invertible, the inverse A^{-1} exists, and the unique solution of (8) is

$$x = A^{-1}b \tag{9}$$

Remark. *If the rank of a square matrix is less than its dimension, then the matrix is singular. In that case,* *there are infinitely many right-hand sides b that have no solution, and the b that have solutions, have infinitely many solutions.*

Least squares for overdetermined systems

In many applications we have more equations than unknowns. Then A is a *tall* matrix, meaning it has more rows than columns. If $A \in \mathbb{R}^{m \times n}$ with $m > n$, the equation

$$Ax = b$$

usually cannot be satisfied exactly for an arbitrary $b \in \mathbb{R}^m$.

In that case, instead of solving $Ax = b$ exactly, we look for the vector x that makes the residual

$$r = Ax - b$$

as small as possible. The standard choice is to minimize the sum of the squares of the residual components, which leads to the *least-squares problem*

$$\min_x \|Ax - b\|^2 \tag{10}$$

Here $\|v\|^2 = v'v$ denotes the square of the Euclidean norm.

To find the minimizer, expand the objective function

$$\|Ax - b\|^2 = (Ax - b)'(Ax - b)$$

Differentiating with respect to x and setting the derivative equal to zero gives the *normal equations*

$$A'Ax = A'b \tag{11}$$

If the columns of A are linearly independent, then $A'A$ is invertible. In that case the least-squares problem (10) has a unique solution,

$$x_{\text{ls}} = (A'A)^{-1}A'b \tag{12}$$

The matrix

$$A^{\dagger} = (A'A)^{-1}A' \tag{13}$$

is called the *pseudo-inverse* of A for this full-column-rank case, so (12) may be written as

$$x_{\text{ls}} = A^{\dagger}b$$

Remark. *The least-squares solution makes the residual vector orthogonal to every column of A. This orthogonality condition is exactly the normal equation* (11).

Example 3.1 (Fitting a straight line to data)**.** Suppose we are given p data points, denoted by (x_i, y_i) for $1 \le i \le p$, and we want to fit a straight line

$$y_i \approx mx_i + b \qquad 1 \le i \le p \tag{14}$$

where m is the slope and b is the intercept.

For each data point, define the residual

$$r_i = y_i - (mx_i + b) \tag{15}$$

Collect these residuals into the vector

$$r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix} \tag{16}$$

Then

$$\sum_{i=1}^{p} r_i^2 = r'r = \|r\|^2 \tag{17}$$

So the least-squares fit chooses m and b to minimize $\|r\|^2$.

To write this in matrix form, define the parameter vector

$$\theta = \begin{bmatrix} m \\ b \end{bmatrix}$$

To avoid overloading notation, we write the vector of measured outputs as y rather than b. Then the data matrix is

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_p & 1 \end{bmatrix} \tag{18}$$

and the data vector is

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \tag{19}$$

Then the residual vector is

$$r = A\theta - y \tag{20}$$

Therefore the line-fitting problem becomes

$$\min_{\theta} \|A\theta - y\|^2 = \min_{\theta} \|r\|^2 \tag{21}$$

This is exactly a least-squares problem of the form (10).

If the x_i are not all equal, the two columns of A are linearly independent, so the least-squares fit is unique and is given by

$$\theta_{\text{ls}} = (A'A)^{-1}A'y \tag{22}$$

Thus the best-fit slope and intercept are the two entries of θ_{ls}.

Figure 4 shows a simple set of data points and the corresponding least-squares line.

When you absolutely, positively have to see all the solutions

So far we have emphasized the case in which a system has a unique solution. If the matrix A is square but singular, then A^{-1} does not exist. In that case there are solutions to $Ax = b$ for some right-hand sides b, but not for all b, and whenever a solution does exist it will not be unique. The same nonuniqueness appears in least squares when the columns of A are not linearly independent.

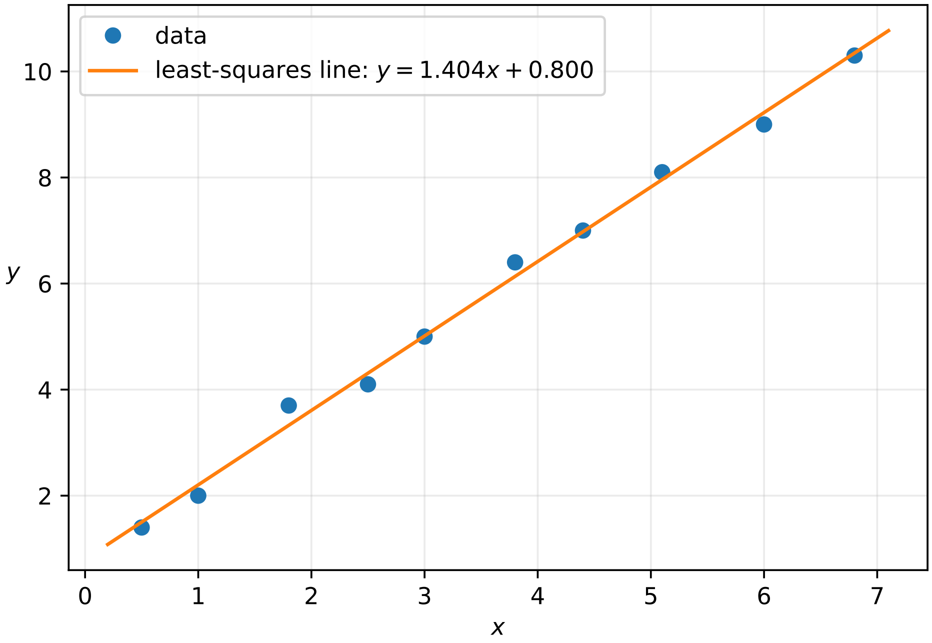

Figure 4: Ten data points (x_i, y_i) shown as circles, together with the least-squares line $y = mx + b$.

To describe *all* solutions, we also solve the homogeneous problem

$$Ax = 0$$

If x_p is one particular solution of $Ax = b$, then every exact solution has the form

$$x = x_p + x_h \qquad Ax_h = 0$$

Similarly, if x_{ls} is one least-squares solution of

$$\min_x \|Ax - b\|^2$$

then every least-squares solution has the form

$$x = x_{\text{ls}} + x_h \qquad Ax_h = 0$$

This is because adding a vector x_h with $Ax_h = 0$ does not change the residual

$$A(x_{\text{ls}} + x_h) - b = Ax_{\text{ls}} - b$$

This is very similar to what happens later in ordinary differential equations. There, one finds a particular solution and then adds the solution of the corresponding homogeneous differential equation to generate all solutions. In linear algebra the corresponding homogeneous problem is again $Ax = 0$, meaning the same matrix A but with the right-hand side set equal to zero. In ordinary differential equations one usually also imposes boundary or initial conditions, and those additional conditions select a unique solution. In linear algebra, adding extra linear conditions simply means adding more equations, which changes the matrix by adding more rows.

Example 3.2 (A singular square system with infinitely many exact solutions)**.** Consider

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Then $Ax = b$ is the system

$$x_1 - x_2 = -1 \qquad -x_1 + x_2 = 1$$

These two equations are the same condition written twice, so there are infinitely many exact solutions. Solving for x_1 gives

$$x_1 = x_2 - 1$$

Hence one particular solution is

$$x_p = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

The corresponding homogeneous problem is

$$Ax = 0$$

which becomes

$$x_1 - x_2 = 0$$

So every homogeneous solution has the form

$$x_h = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \alpha \in \mathbb{R}$$

Therefore every exact solution of $Ax = b$ is obtained by adding the homogeneous solution to the particular solution

$$x = x_p + x_h = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \alpha \in \mathbb{R}$$

The minimum-norm solution is the point on this line closest to the origin. Minimizing $x_1^2 + x_2^2$ subject to $x_1 - x_2 = -1$ gives

$$x_{\min} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

Figure 5 shows the full line of exact solutions, the homogeneous line $Ax_h = 0$, and the minimum-norm solution.

Example 3.3 (A least-squares problem with infinitely many minimizers)**.** Return to the line-fitting problem, but now set $x_i = 4$ for every data point while keeping the same y_i values used in Figure 4. Then

$$A = \begin{bmatrix} 4 & 1 \\ 4 & 1 \\ \vdots & \vdots \\ 4 & 1 \end{bmatrix}$$

The two columns are not linearly independent, since the first column is four times the second. Therefore the least-squares minimizer is not unique.

If

$$\bar{y} = \frac{1}{p} \sum_{i=1}^{p} y_i$$

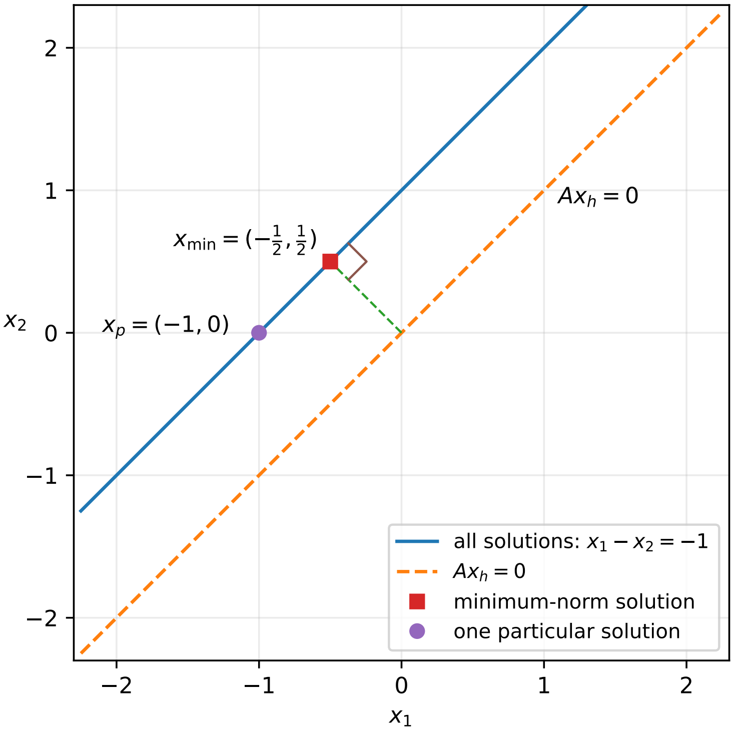

Figure 5: All exact solutions of the singular system $Ax = b$ lie on the line $x_1 - x_2 = -1$. The dashed line through the origin is the homogeneous solution set $Ax_h = 0$. The minimum-norm solution is the point on the affine solution line closest to the origin, and with identical axis scales the perpendicular geometry is easy to see.

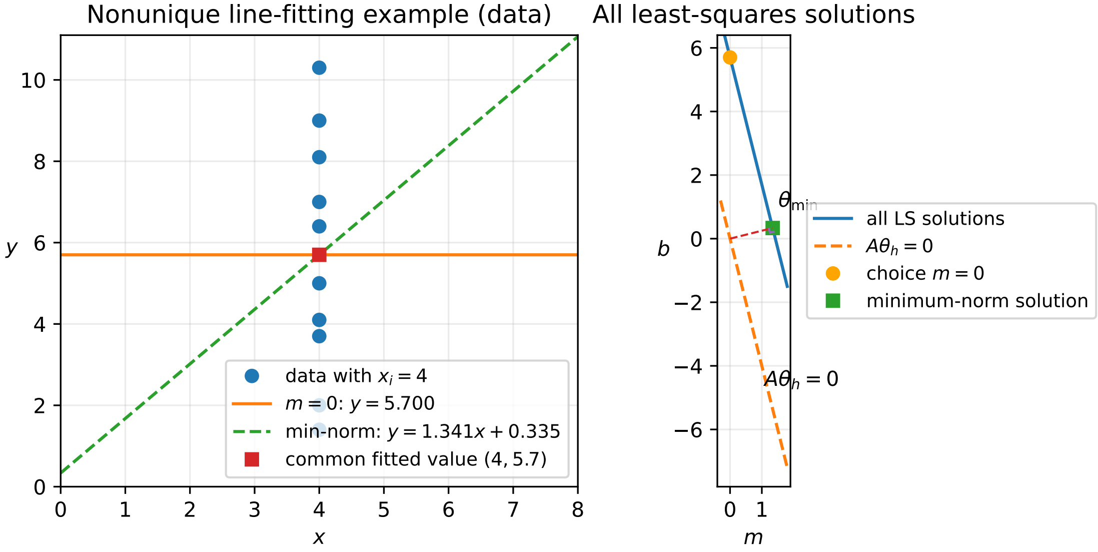

Figure 6: Left: the same ten y_i values are plotted against $x_i = 4$, together with the fitted line obtained from the choice $m = 0$ and the fitted line obtained from the minimum-norm solution. Both lines give exactly the same least-squares objective value because they agree at $x = 4$. Right: all least-squares solutions for the modified line-fitting example satisfy $4m + b = 5.7$. The dashed line is the corresponding homogeneous set $A\theta_h = 0$, and the horizontal and vertical axis scales are identical, so the minimum-norm solution is visibly the point on this line closest to the origin.

then every fitted value is equal to $4m + b$, so the least-squares objective depends only on that combination. It is minimized when

$$4m + b = \bar{y}$$

For the data in Figure 4, one finds

$$\bar{y} = 5.7$$

Therefore the set of all least-squares solutions is the straight line

$$4m + b = 5.7$$

One convenient solution is obtained by setting $m = 0$, which gives

$$\theta_{m=0} = \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 5.7 \end{bmatrix}$$

Among all these minimizers, the minimum-norm solution is the point on the line $4m + b = 5.7$ that is closest to the origin. Minimizing $m^2 + b^2$ subject to $4m + b = 5.7$ gives

$$\theta_{\min} = \begin{bmatrix} m \\ b \end{bmatrix} = \frac{5.7}{17} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.341 \\ 0.335 \end{bmatrix}$$

Coding. In MATLAB, the minimum-norm least-squares solution for this rank-deficient problem is obtained with `lsqminnorm(A,y)` or `pinv(A)*y`. The backslash command `A\y` returns a least-squares solution, but it is not guaranteed to be the minimum-norm one when A is rank deficient.

In Python, the corresponding computation is `numpy.linalg.pinv(A) @ y`. SciPy provides the analogous call `scipy.linalg.pinv(A) @ y`. Figure 6 shows the full line of least-squares solutions in the (m, b) plane and compares the fitted lines corresponding to the choice $m = 0$ and the minimum-norm solution.

4 Eigenvalues and eigenvectors

For a real-valued, square matrix $A \in \mathbb{R}^{n\times n}$, a nonzero vector $v \in \mathbb{R}^n$ is called an *eigenvector* of A with corresponding *eigenvalue* $\lambda \in \mathbb{C}$ if

$$Av = \lambda v, \qquad v \neq 0 \tag{23}$$

Key facts about eigenvalues and eigenvectors

- **An $n \times n$ matrix has n eigenvalues** (counted with multiplicity), but they may be complex-valued.
- **If A is symmetric, then all eigenvalues are real.** To see this, suppose A is symmetric and λ is an eigenvalue with eigenvector v (which may be complex). Then

  $$Av = \lambda v$$

  Taking the complex conjugate of both sides and using $\overline{A} = A$ (since A is real),

  $$A\overline{v} = \overline{\lambda}\overline{v}$$

  So $\overline{v}$ is an eigenvector with eigenvalue $\overline{\lambda}$. Now compute

  $$(\overline{v}, Av) = (\overline{v}, \lambda v) = \lambda(\overline{v}, v)$$

  and also

  $$(\overline{v}, Av) = (A^T\overline{v}, v) = (A\overline{v}, v) = (\overline{\lambda}\overline{v}, v) = \overline{\lambda}(\overline{v}, v)$$

  Since $(\overline{v}, v) = \sum_i |\overline{v}_i|^2 > 0$ (the eigenvector is nonzero), we can divide to get $\lambda = \overline{\lambda}$, meaning λ is real.
- **If A is symmetric, eigenvectors corresponding to different eigenvalues are orthogonal.** Let v_1 and v_2 be eigenvectors with distinct eigenvalues $\lambda_1 \neq \lambda_2$. Then

  $$Av_1 = \lambda_1 v_1 \qquad Av_2 = \lambda_2 v_2$$

  Computing the inner product (v_2, Av_1) in two ways

  $$(v_2, Av_1) = (v_2, \lambda_1 v_1) = \lambda_1(v_2, v_1)$$

  and using $A^T = A$,

  $$(v_2, Av_1) = (A^T v_2, v_1) = (Av_2, v_1) = (\lambda_2 v_2, v_1) = \lambda_2(v_2, v_1)$$

  Subtracting these results gives

  $$0 = (\lambda_1 - \lambda_2)(v_2, v_1)$$

  Since $\lambda_1 \neq \lambda_2$, we have $(v_2, v_1) = 0$, i.e., $v_1 \perp v_2$.
- **Even if A is symmetric with repeated eigenvalues, we can always choose n orthogonal eigenvectors.** When an eigenvalue is repeated, the corresponding eigenspace may have dimension greater than 1, giving us freedom in which eigenvectors to select. Orthogonalizing within that subspace ensures orthogonality.

Examples

Example 4.1 (A 2×2 real matrix with complex eigenvalues)**.** Consider the rotation matrix

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

The characteristic equation is $\det(A - \lambda I) = \lambda^2 + 1 = 0$, so the eigenvalues are $\lambda = \pm i$. These are complex, and no real eigenvectors exist (the matrix represents a 90-degree rotation).

Example 4.2 (A symmetric matrix with real, distinct eigenvalues)**.** Consider

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

This is symmetric, so eigenvalues are real. The characteristic equation is $\det(A - \lambda I) = (3 - \lambda)(1 - \lambda) - 1 = \lambda^2 - 4\lambda + 2 = 0$. Using the quadratic formula, $\lambda = 2 \pm \sqrt{2}$.

For $\lambda_1 = 2 + \sqrt{2}$, solving $(A - \lambda_1 I)v = 0$ gives an eigenvector $v_1 = [1, -1 - \sqrt{2}]^T$ (or any scalar multiple).

For $\lambda_2 = 2 - \sqrt{2}$, solving $(A - \lambda_2 I)v = 0$ gives $v_2 = [1, -1 + \sqrt{2}]^T$.

We can verify orthogonality

$$(v_1, v_2) = 1 \cdot 1 + (-1 - \sqrt{2})(-1 + \sqrt{2}) = 1 + (1 - 2) = 0$$

Eigenvector orthogonality with repeated eigenvalues

When a symmetric matrix has repeated eigenvalues, eigenvectors can still be chosen to be mutually orthogonal. See Exercise 11 for an illustration.

Exercises

Exercise 1 (Trigonometric and hyperbolic functions and their zeros).

(a) Make a plot of $\cosh x$ and $\sinh x$ for real-valued x.

(b) Using Euler's formula, derive the two relationships stated in the handout valid for all complex-valued z

$$\cosh iz = \cos z \qquad \sinh iz = i \sin z$$

(c) In the complex plane, sketch the locations that you already know for the zeros of $\sin z$ and $\cos z$. Note that the real zeros that you know are *all* the zeros in the complex plane.

(d) Using your derived relationship between the trigonometric and hyperbolic function, sketch the locations of all the zeros of $\cosh z$ and $\sinh z$. So now you know that the hyperbolic functions also have infinitely many zeros, but we have to consider the complex plane to locate them.

(e) Bonus question. So what about the exponential itself, which we are using to build the trigonmetric and hyperbolic functions. Does e^z have zeros in the complex plane?

Exercise 2 (Matrix inverse and linearly independent columns).

The humble 2×2 matrix has an outsized importance because it is the largest matrix that I would ever ask you to manipulate by hand on an exam. So consider the 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

(a) For what values of a_{ij} entries is this matrix invertible?

(b) If the matrix satisfies this condition, what is the matrix inverse A^{-1}? Remember how to get this result quickly and correctly because you may need it on the midterm or final exam.

(c) What does it mean for a matrix with only two columns that those two columns are linearly independent? Show that if the columns are not linearly independent, then you do not satisfy the condition stated in (a) for the matrix to be invertible.

Exercise 3 (Least-squares estimation of activation energy[1]).

Assume you have measured a rate constant, k, at several different temperatures, T, and wish to find the activation energy (divided by the gas constant), E/R, and the preexponential factor, k_0, in the Arrhenius model

$$k = k_0 e^{-E/RT} \tag{24}$$

The data are shown in Figure 7 and listed here.

T (K)	300	325	350	375	400	425	450	475	500
k	2.03	2.02	2.09	2.19	2.21	2.10	2.22	2.19	2.22

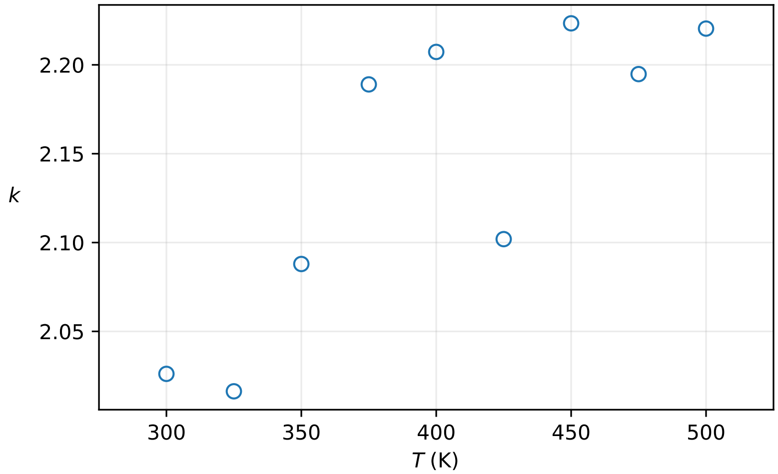

Figure 7: Measured rate constant at several temperatures.

(a) Take logarithms of (24) and write a model that is linear in the parameters $\ln(k_0)$ and E/R. Summarize the data and model with the linear algebra problem

$$Ax = b$$

in which x contains the parameters of the least-squares problem

$$x = \begin{bmatrix} \ln(k_0) \\ E/R \end{bmatrix}$$

What are A and b for this problem?

(b) Find the least-squares fit to the data. What are your least-squares estimates of $\ln(k_0)$ and E/R?

(c) Is your answer unique? How do you know?

(d) Plot the data and least-squares fit in the original variables k versus T. Do you have a good fit to the data?

Exercise 4 (Convergence of series by the ratio test).

Use the ratio test to determine whether each of the following series converges or diverges.

(a)

$$\sum_{n=1}^{\infty} \frac{n}{2^n}$$

(b)

$$\sum_{n=0}^{\infty} \frac{3^n}{n!}$$

---

[1] See also (Graham and Rawlings, 2022, Exercise 1.22).

(c)

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}$$

Exercise 5 (Radius of convergence of a power series).

A power series $\sum_{n=0}^{\infty} c_n x^n$ converges for all $|x| < \rho$, where the radius of convergence $\rho = 1/R$ and $R = \lim_{n\to\infty} |c_{n+1}/c_n|$ is the ratio-test limit for the series with $a_n = c_n x^n$.

(a) Find the radius of convergence of the power series

$$\sum_{n=0}^{\infty} \frac{x^n}{n+1}$$

(b) Show that the Taylor series for e^x,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

converges for all real (and complex) x, i.e. show that $\rho = \infty$.

Exercise 6 (Algebra and polar form of complex numbers).

Let $z_1 = 3 + 4i$ and $z_2 = 1 - 2i$.

(a) Compute $z_1 + z_2$, $z_1 z_2$, and z_1/z_2 by performing the algebra in rectangular form $a + ib$.

(b) Write z_1 and z_2 in polar (exponential) form $re^{i\theta}$. Find r and θ for each.

(c) Use the polar forms to recompute $z_1 z_2$ and z_1/z_2, and verify that your answers are consistent with part (a).

Exercise 7 (Euler's formula and De Moivre's theorem).

(a) Use Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ to derive the double-angle identities

$$\cos 2\theta = \cos^2\theta - \sin^2\theta \qquad \sin 2\theta = 2\sin\theta\cos\theta$$

Hint: compute $e^{i(2\theta)} = (e^{i\theta})^2$ *and separate real and imaginary parts.*

(b) Express $\cos\theta$ and $\sin\theta$ in terms of complex exponentials. Use these expressions to show

$$\cos^2\theta + \sin^2\theta = 1$$

Exercise 8 (Solving $Ax = b$ and linear independence).

Consider the linear system

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}, \qquad b = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

(a) Compute $\det A$ and use it to write down A^{-1} for this 2×2 matrix.

(b) Find the unique solution $x = A^{-1}b$.

(c) Now consider the matrix

$$B = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

Are the columns of B linearly independent? What is $\det B$, and what does this imply about solving $Bx = c$ for a general right-hand side c?

Exercise 9 (Least-squares fit to noisy linear data).

Suppose you measure the response y at four values of an input t and obtain the data

t	0	1	2	3
y	1.1	2.9	5.2	6.8

You believe the underlying model is linear: $y = \beta_0 + \beta_1 t$.

(a) Write the over-determined system $Ax = b$ where $x = [\beta_0,\ \beta_1]^T$. What are A and b?

(b) Write down the normal equations $A^T Ax = A^T b$ and compute the matrices $A^T A$ and $A^T b$.

(c) Solve the normal equations to find the least-squares estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ with your favorite package (matlab, excel, python), and then show that you get the same solution by hand.

Exercise 10 (Spectroscopy and least squares—ideal mixture).

The basics of spectroscopy are to study molecules by exposing them to various frequencies of electromagnetic radiation and observing the emission or absorption of the sample. Here we look at the basic mathematical problem of inferring the composition of a sample based on its measured spectrum and the spectrum of a standard comprising samples of known composition.

We consider a ternary mixture of three different molecular species, denoted A, B and C. Let the mixture's molar fraction be denoted x_A, x_B, x_C. We wish to estimate the composition of the mixture from its spectroscopic measurement shown in Figure 8.

To do this estimation, we will first measure each of the pure components. Denote the absorption/emission signal strength of the pure components as a function of radiation frequency (or wavenumber) by $s_A(\omega_i), s_B(\omega_i), s_C(\omega_i)$, with $i = 1, 2, \ldots, f$ where f is the number of frequencies at which we collect the spectrum. Typically we measure spectra at hundreds of frequencies. Figure 9 shows the absorption spectra of the pure samples of A, B, and C.

Ideal mixture. We start the analysis with a simple model for the mixture, the ideal mixture assumption. For an ideal mixture with composition (x_A, x_B, x_C) we assume the spectrum of a mixture s_m is given by the linear combination of the pure component spectra weighted by the corresponding mole fraction in the mixture

$$s_m(\omega_i) = x_A s_A(\omega_i) + x_B s_B(\omega_i) + x_C s_C(\omega_i), \qquad i = 1, 2, \ldots, f \tag{25}$$

(a) As we did for fitting the straight line example, define a y vector, X matrix, and θ vector so that the ideal mixture equation (25) can be expressed with one equation for all frequencies

$$y = X\theta \tag{26}$$

What are y, X and θ for your composition estimation problem. What are the dimensions of these vectors and matrices?

(b) We do not expect to be able to solve this linear algebra problem. Why not? How many equations do you have in (26), and how many unknowns do you have?

(c) So instead we are going to use least squares and solve the following problem

$$\min_{\theta} \|y - X\theta\|^2 \tag{27}$$

What is the formula for the solution, $\hat{\theta}$, to the least squares problem (27)?

(d) What restriction must hold for your least-squares formula to be valid? Looking at Figure 9, do you think this restriction holds in this case? Why or why not?

Exercise 11 (Symmetric matrix with repeated eigenvalues).

Consider the symmetric 2×2 matrix

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

(a) What are the eigenvalues of A? Are they repeated?

(b) Find two linearly independent eigenvectors v_1 and v_2 corresponding to this repeated eigenvalue.

(c) Show that the eigenvectors you found are orthogonal, and explain why this makes sense for a symmetric matrix.

References

M. D. Graham and J. B. Rawlings. *Modeling and Analysis Principles for Chemical and Biological Engineers*. Nob Hill Publishing, Santa Barbara, CA, 2nd, paperback edition, 2022. 560 pages, ISBN 978-0-9759377-6-1.

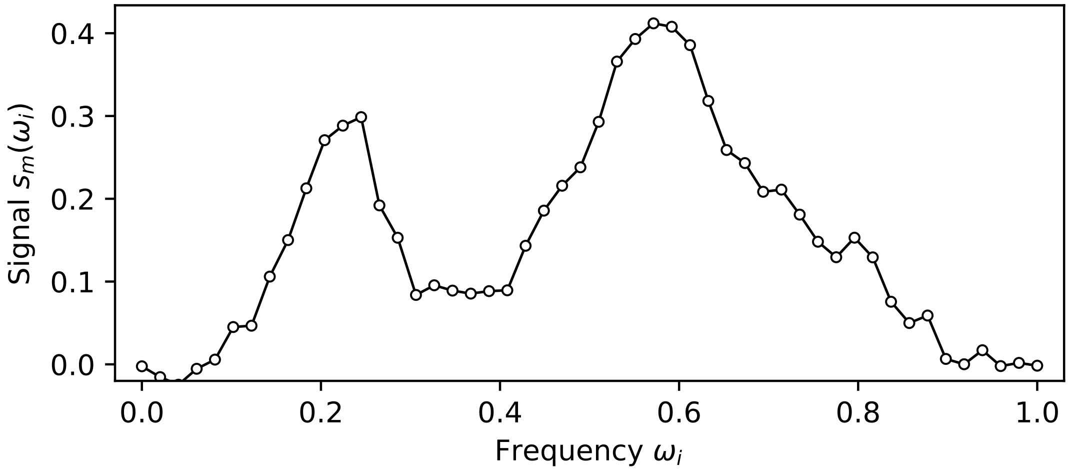

Figure 8: Measured spectrum s_m for a mixture with unknown composition (x_A, x_B, x_C).

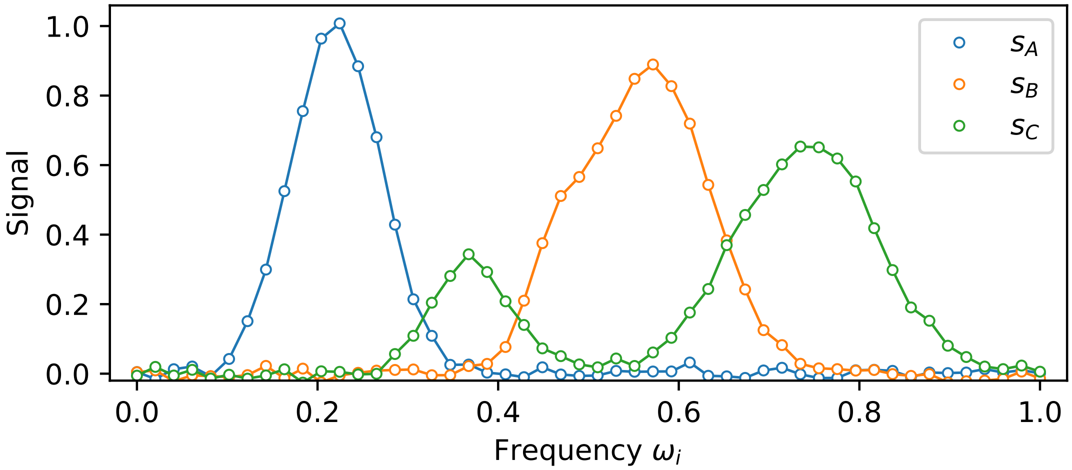

Figure 9: Measured absorption/emission spectra of pure components s_A, s_B, and s_C.