General optimisation problems

Use calculus, common sense or a graphic argument to answer the following optimization questions (give an optimal solution and the optimal objective value if it exists; if not, argue why):

1. Minimize $x^2 + 1$ for $x \in \mathbb{R}$;
2. Maximize $x + y$ for $x, y \in [-1, 1]$;
3. Maximize $x + y$ for $x, y \in (-1, 1)$;
4. `Optional:` Minimize $e^{x^2+y^2+z^2}$ for $2x^2 + 3y^2 + 4z^2 < 1$; How about replacing Max by Min and < 1 by ≤ 1?
5. Maximize $x^2 + y^2$ for $|x| + |y| = 1$;
6. Maximize $y - x$ for $x^2 - y^2 \geq 1$, $x \geq 0$. HINT: $x^2 - y^2 = 1$ is a hyperbola.

Linear programming

1. **Formulate as a LP, but do not try to solve:** A toy factory makes three types of wooden dolls: Grumpy, Sleepy, and Bashful. The manufacturing process uses three different machines A,B,C. Each Grumpy is processed for one hour each on machines A and C and three hours on B. Sleepy takes two hours on machine A and four hours on C. Bashful takes an hour on machine A and two hours on machine B. Because of maintenance schedules, machine A is available for 40 hours a week, B for 50 and C for 45. The profit per item is £4, £2, £5 respectively for the three types. How many of each type should the factory make per week in order to maximize the total profit?

   Suppose now there is an extra condition that exactly 360 toys are to be manufactured each week. Use this to reduce the number of variables, so the problem can be solved graphically. Do the graphs and show that such a problem is unfeasible: the feasible set is empty.

2. **Solve graphically a LP:** A small-scale farmer has 100 acres of land. He can use it as pasture for cows, at a profit of £10 per acre per year, or as arable land for growing crops, at a profit of £15 per acre per year; or he can leave it fallow at zero profit. Pasture requires 30 hours work per acre per year, arable needs 60 hours work per acre per year. The farmer wants to maximize profits, and is willing to work 4200 hours per year.

   (a) How should the land be divided between pasture and arable?
   (b) How much extra profit will the farmer get, if he acquires an extra acre of land, the rest of the data being the same?
   (c) What if he still has 100 acres, but the profit from pasture changes to £7 per acre per year?

3. **Optional: Formulate as LP, do not solve.** A Manufacturer of herbal medicine distills three essential medicine Components out of three different plant Extracts. The three Components are combined to form two different mixtures, called Remedy 1 and Remedy 2. The Manufacturer has to decide on the amounts of Extracts to buy and Remedies to produce, in order to maximize his gains. Formulate the task as a LP, given that:

[1] If you spot any bugs, please let me know.

- To produce one gram of Remedy 1, one needs .25g., .35g. and .15g. of Components 1,2, and 3, respectively. The rest is alcohol, which is available free and unlimited (what a world would that be!). Remedy 1 sells for £10 per gram and the manufacturer cannot sell more than 100 grams of it.
- To produce one gram of Remedy 2, one needs .20g., .10g. and .30g. of Components 1,2, and 3, respectively. The rest is alcohol. Remedy 2 sells for £13 per gram and the manufacturer cannot sell more than 130 grams.
- One gram of Extract 1 contains .20g., .15g., and .25g. of Components 1,2, and 3, respectively, costs £3, and is available in unlimited quantities.
- One gram of Extract 2 contains .30g., .30g., and 0g. of Components 1,2, and 3, respectively, costs £4, and no more than 1000 grams are available on the market.
- One gram of Extract 3 contains .10g., .15g., and .45g. of Components 1,2, and 3, respectively, costs £5, and no more than 500 grams are available.
- Distilling the Components out of Extracts and mixing the Remedies does not involve any additional costs.

4. **Optional: Formulate the following optimal assignment task as a LP.** A group of four Universities in the South-West of England has to form a student team for a quiz show. The conditions of the show are as follows. One team member is to be selected from each University. The team member must be enrolled in one of the following single honours Degree Programmes: Mathematics, Physics, Chemistry, or Biology. No two team members are allowed to come from Degree Programmes in the same subject. Each Degree Programme in each University presents one candidate, and on the basis of a series of tests, each candidate is assigned an Erudition coefficient e_{ij}, where i marks the University and j marks the Degree Programme that the candidate comes from. The objective now is to maximise the overall strength of the team, i.e., the sum of the Erudition coefficients of chosen candidates.

   HINT: Introduce the unknowns $x_{ij} = \begin{cases} 1 & \text{if the corresponding candidate is chosen,} \\ 0 & \text{otherwise.} \end{cases}$

Optional section: Some review problems in linear algebra

This section is a bit lengthy. Its purpose, however, is to bring back the key notions and techniques from linear algebra that this course needs. The problems are followed by some brief theoretical overview that may be useful for solving them.

Notation – important to avoid confusion! The vector-notation $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is identified with the matrix *column*-vector notation

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}.$$

The same vector can be, of course, represented by the single-column matrix, or the *row*-vector

$$\boldsymbol{x}^T = [x_1 \; x_2 \; \ldots \; x_n],$$

the superscript ^{T} to be read as *transpose*. There are no comas separating the components in the matrix notation and the brackets are square. All this is just a matter of bookkeeping: the matrix notation is undoubtedly convenient, but one can certainly jot down the components of the vector $\boldsymbol{x}$ either one under the other – as a column-matrix $\boldsymbol{x}$ – or one after the other – as the row-matrix $\boldsymbol{x}^T$. On the other hand, the rules of matrix multiplication are unambiguous. So, the dot product $\boldsymbol{x} \cdot \boldsymbol{y}$ of two vectors can be represented using the matrix notation as

$$\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^T \boldsymbol{y} = \boldsymbol{y}^T \boldsymbol{x} = \boldsymbol{y} \cdot \boldsymbol{x}.$$

On the other hand, if one looks at the expression $\boldsymbol{x}\boldsymbol{y}^T$ (a column-matrix, multiplied by a row-matrix, then the result, according to the rules of matrix multiplication, is a $n \times n$ matrix, whose ij-components equal $x_i y_j$.

Problems:

1. For a vector $\boldsymbol{u} = (1, 2, 3)$ find the products $\boldsymbol{u}\,\boldsymbol{u}^T$ and $\boldsymbol{u}^T\,\boldsymbol{u}$.

   Argue that no matter what the size of a matrix A, the product AA^T (and thus A^TA) is always defined. If A is $m \times n$, what are the sizes of A^TA and AA^T?

   For a pair of matrices
   $$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 0 & 5 \end{bmatrix},$$
   find the products AA^T, A^TA, BB^T, B^TB, AB, BA, A^2, B^2 or state that they are not defined.

2. Suppose A is a 4×5 matrix. Find a 4×4 matrix B, such that the product BA is a matrix, whose third row is the sum of the first and the third rows of A, while the rest of the rows are the same as the corresponding rows of A.

   Find a 5×5 matrix C, such that the product AC equals a matrix, obtained from A by swapping the second and the fifth columns.

3. Find all the solutions (if any) of the following systems of linear equations by Gaussian elimination:

$$\left\{\begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & + & x_3 & = & 9 \\ 2x_1 & - & x_2 & + & 2x_3 & = & 6 \\ x_1 & - & x_2 & + & 2x_3 & = & 5 \end{array}\right. , \quad \left\{\begin{array}{rcrcrcrcr} 2x_1 & - & x_2 & + & x_3 & + & x_4 & = & 6 \\ x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 8 \end{array}\right. ,$$

$$\left\{\begin{array}{rcrcrcr} x_1 & - & x_2 & + & x_3 & = & 5 \\ x_1 & + & 2x_2 & - & x_3 & = & 3 \\ 2x_1 & + & x_2 & & & = & 7 \end{array}\right. .$$

4. Let

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & 2 \\ 1 & -1 & 2 \end{bmatrix}.$$

Argue that *any* number of systems of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ with the *same* matrix A, the unknown $\boldsymbol{x} = (x_1, x_2, x_3)$, and *different* right-hand sides $\boldsymbol{b}^1$, $\boldsymbol{b}^2, \ldots \in \mathbb{R}^3$, can be solved simultaneously as the result of the following procedure.

Write an "extended" matrix

$$\tilde{A} = [A \, |\boldsymbol{b}^1 \, \boldsymbol{b}^2 \ldots]$$

and apply the Gaussian elimination algorithm to it (certainly (if possible) pivot only the (non-zero) entries positioned to the left of the vertical bar, e.g. those on the main diagonal). Thereupon, the columns to the left of the vertical bar will contain the solutions $\boldsymbol{x}^1$, $\boldsymbol{x}^2, \ldots$ corresponding to the right-hand sides $\boldsymbol{b}^1$, $\boldsymbol{b}^2, \ldots$.

Argue that in fact, by doing this you are solving a linear equation $AX = B$ for the unknown *matrix* X, whose columns are the vectors $\boldsymbol{x}^1$, $\boldsymbol{x}^2, \ldots$, the right-hand side being a matrix B, whose columns are the vectors $\boldsymbol{b}^1$, $\boldsymbol{b}^2, \ldots$. In particular, for the matrix A above, if $\boldsymbol{b}^1 = (1,0,0)$, $\boldsymbol{b}^2 = (0,1,0)$, $b^3 = (0,0,1)$, the unknown matrix X is the inverse A^{-1} of A.

Hence, find the inverse of A by the Gauss-Jordan method. Find the determinant of A.

5. Find $\texttt{rank}\, A$ for

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

6. Prove $(AB)^T = B^T A^T$, $(AB)^{-1} = B^{-1}A^{-1}$, the latter for square invertible matrices.

7. Let A be a square $n \times n$ matrix. Present an argument that the rows of A are linearly dependent, if and only if the columns of A (the rows of A^T) are also linearly dependent (you may use a geometric argument, or use determinants, or talk about an inverse).

Linear algebra: a review

This section is to review some basic notions of linear algebra to be used throughout this course. It gives a compendium of remarks on notation in the above problems, as well as some brief theoretical statements. However, it may be necessary for you to consult the Linear algebra section of your last year notes for Core Mathematics B or one of the texts, recommended for that course.

Notation and theory

- Matrices are denoted by capital letters. An $m \times n$ matrix is a table of real numbers with m rows and n columns. A table is put into square brackets, e.g.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

  A matrix, consisting of one column only is called a column-vector, or a *vector.* A matrix, consisting of a single row is called a row-vector, or a *covector.* Vectors are usually denoted by **bold** lowercase letters. For a vector $\boldsymbol{x}$ with n components, one writes $\boldsymbol{x} \in \mathbb{R}^n$. Components of $x \in \mathbb{R}^n$ are marked by subscripts, i.e.

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix}.$$

  Thus, an $m \times n$ matrix A can be written as $A = [\boldsymbol{a}^1\, \boldsymbol{a}^2, \ldots \boldsymbol{a}^n]$, where $\boldsymbol{a}^j$, $j = 1, 2, \ldots, n$ are the columns of A, each one being an m-vector: $\boldsymbol{a}^j \in \mathbb{R}^m$.

- A transpose A^T of an $m \times n$ matrix A is an $n \times m$ matrix, whose respective columns are the rows of A. For instance, for the above A and $\boldsymbol{x}$,

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad \boldsymbol{x}^T = [x_1\, x_2 \ldots x_n].$$

  In particular, for any matrix A, $[A^T]^T = A$; a transpose of a vector is a covector, a transpose of a covector is a vector.

- To save some paper, to denote column-vectors we will also use the notation $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, round brackets and commas indicating that this is just an economical way to write a column-vector, but not a row vector. Thus, $(x_1, x_2, \ldots, x_n) = [x_1\, x_2 \ldots, x_n]^T$.

- An *identity* or *unit* matrix of size m (often denoted as Id, I, I_m, whatever happens to be unambiguous) is a square $m \times m$ matrix, such that the *main diagonal* elements a_{ii}, $i = 1, \ldots, m$ equal 1, while the rest of the elements a_{ij}, $i \neq j$, $i, j = 1, \ldots, m$, equal 0. One sometimes writes $I_m = \text{diag}(1, \ldots, 1)$, emphasising that this is a *diagonal* matrix.

- For any matrix A, let us define three types of *elementary row operations* (alias eros):

  1. interchange any pair of rows;
  2. multiply a row by a nonzero real number;
  3. add a multiple of one row to another row.

  Note: a single ero leaves the rows, which are not involved unchanged. Also, each ero is invertible, i.e. can be undone by some other ero. E.g. multiplying a row by 2 can be further undone by multiplying the same row by .5, etc. A single ero, applied to a matrix results in a new matrix of the same size. The result of fulfilling a sequence of eros in general depends on the order in which these eros are to be performed.

  Also note that applying a single ero to a matrix A can be described as the result of multiplying this matrix by some specific matrix B, corresponding to this specific ero, from the left, i.e.

computing BA. For instance, multiplying the second row of an $m \times n$ matrix A by 2 boils down to taking the matrix product BA of a square $m \times m$ matrix $B = \mathrm{diag}(1, 2, 1, \ldots, 1)$ with A (so, B is just an identity matrix I_m, whose element at the intersection of the second row and the second column has been changed from 1 to 2).

- For a matrix, *pivoting* a *non-zero* element a_{ij} (or pivoting *about* or *around* the element a_{ij}), located at the intersection of the ith row and the jth column, referred to as a *pivot row* and a *pivot column* (another common notation for a_{ij} is a^i_j, to easier distinguish the row, upper and the column, lower indices) is a sequence of the following eros:

  1. divide the *pivot row* by a_{ij}, whereupon the latter becomes one;
  2. by adding an appropriate multiple of the pivot row to all other rows, ensure that all the rest of the entries in the *pivot column* become zero.

  Thus, a pivot transforms the pivot column into a column of the unit matrix.

- The Gaussian elimination algorithm (alias the Gauss-Jordan method) of solving a system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$, where A is $m \times n$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, consists in the following:

  1. write down an "extended" matrix
  $$\tilde{A} = [A \,|\boldsymbol{b}]$$
  (note, $\tilde{A}$ is $m \times (n+1)$; the first n columns can be labeled by $x_1, x_2, \ldots, x_n$, as the entries $\tilde{a}_{ij} = a_{ij}$ in these columns are the coefficients, multiplying x_j in the ith equation, $i = 1, \ldots, m$, $j = 1, \ldots, n$; the vertical bar has no mathematical meaning and is drawn exclusively for convenience, to separate the last column of $\tilde{A}$, corresponding to the equations' right-hand sides);
  2. perform *any* succession of pivots, all the pivot elements sitting in different rows, *to the left* of the vertical bar;
  3. read out the solution (see below).

  Note: the most convenient sequence of elements to pivot are those, sitting on the *main diagonal* of A, i.e those whose row and column indices are the same, like a_{11}. However, this can be done only if these elements are non-zero (in due time throughout the fulfillment of the algorithm). The pivoting stage ends after either in each row a pivot has been done about some non-zero element in this row, or if after you've done a number of pivots, *all* the elements in yet not pivoted rows to the left of the vertical bar have accidentally become zeroes (which means that the rows of A are linearly dependent). Besides, note that if pivot some element in a specific column, all the rest of the elements in this column will have become zeroes, so one cannot further pivot any other element in this column. Hence, to solve a system of m linear equations, it suffices to complete at most m pivots.

- After completion of the Gaussian elimination procedure, one calls *basic* the components of $\boldsymbol{x}$, marking those columns of A, which contain the positions $_{ij}$, which have been pivoted. The rest (if any) of the components of $\boldsymbol{x}$ are called *free.* Usually one deals with free components only in the case when the number of variables n exceeds the number of equations m. For instance, if there is a system of three equations with four unknonwns (x_1, x_2, x_3, x_4), having done pivots in the corresponding extended matrix about the elements, sitting at positions $_{11}$, $_{22}$, $_{33}$ means that x_1, x_2, x_3 *have been chosen* as basic variables, and have all been expressed in terms of a single free variable x_4. I emphasize *have been chosen*, because one can further express x_4 via x_1 from the first relation, and plug it into the rest of the equations (that is pivot an element sitting at the position $_{14}$), whereupon then the basic variables will be x_2, x_3, x_4, leaving x_1 as a free one.

- In general, the last step of the Gaussian elimination procedure means reading out the expressions for the basic components of $\boldsymbol{x}$ in terms of the right-hand side (where $\boldsymbol{b}$ is transformed after the pivots) and the free components (if any), which can attain *any* values. Let me repeat that in case $n > m$ (usually), the choice of the basic and free components is not unique (e.g. $x_1 + x_2 = 1$ can be solved as $x_1 = 1 - x_2$ or $x_2 = 1 - x_1$, in the former case x_1 being basic and x_2-free, and in the latter case - the other way around.) The only invariant is *the number* of basic components, which equals $\texttt{rank}\, A$; this can actually be taken for the definition of the rank (if $n \geq m$), being also pretty much the only way to practically determine it.

- A *basic* solution corresponds to assigning a zero value to all the free variables, if any (e.g. for a single equation $x_1 + x_2 = 1$, basic solutions are $(x_1, x_2) = (1, 0)$ or $(0, 1)$.)

- On the last step of the elimination algorithm one can conclude, whether the system of linear equations either has a unique solution (which is the case if $m = n = \texttt{rank}\, A$), or infinitely many solutions, or no solutions. If the last possibility takes place, the system is said to be *inconsistent*; pivoting it eventually results in an equation $0 = 1$ (that is a row of zeroes to the left of the vertical bar and a nonzero entry to the right of the bar in the extended matrix).

- A linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if the vector $\boldsymbol{b}$ is a *linear combination* of the columns $\boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^n$ of the matrix A. This is a tautology: indeed, $x_1, x_2, \ldots, x_n$ correspond to the coefficients in this linear combination. A linear combination $A\boldsymbol{x} = x_1\boldsymbol{a}^1 + x_2\boldsymbol{a}^2 + \ldots + x_n\boldsymbol{a}^n$ is *non-trivial* if at least one component of $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ is nonzero (a trivial linear combination always equals zero). Thus, a *homogeneous* system of linear equations $A\boldsymbol{x} = 0$ has a non-zero (non-trivial) solution if and only if the columns of A are *linearly dependent*. In other words, a bunch of vectors is linearly dependent if one of them is a linear combination of the others. Note: more than m vectors in $\mathbb{R}^m$ (with m components) are *always* linearly dependent. E.g. if one has four (or more) vectors in $\mathbb{R}^3$, one of them can be expressed as a linear combination of the others.