Matricies

- Used for graphics, solving systems of linear equations, optimisation, graph representation, machine learning and loss matricies.

Vectors

- One column: magnitude and direction.
- Operations: + and −
  × by a scalar (number), vector or matrix.
- Eg. $\begin{pmatrix}1\\4\end{pmatrix} + \begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}3\\6\end{pmatrix}$ $\begin{pmatrix}2\\3\end{pmatrix} \times 6 = \begin{pmatrix}12\\18\end{pmatrix}$
  ditto for −

Matrix Multiplication

2×2: $\begin{pmatrix}1 & 2\\4 & 3\end{pmatrix} \times \begin{pmatrix}-1 & 2\\3 & 1\end{pmatrix} = \begin{pmatrix}5 & 4\\5 & 11\end{pmatrix}$

$(1\times -1) + (2\times 3) = 5$
$(1\times 2) + (2\times 1) = 4$
$(4\times -1) + (3\times 3) = 5$
$(4\times 2) + (3\times 1) = 11$

3×3: $\begin{pmatrix}1 & 3 & 2\\0 & 3 & 1\\2 & 1 & 3\end{pmatrix} \times \begin{pmatrix}2 & 0 & -1\\1 & 3 & 2\\0 & -1 & 1\end{pmatrix} = \begin{pmatrix}5 & 7 & 7\\3 & 8 & 7\\5 & 0 & 3\end{pmatrix}$

$(1\times 2) + (3\times 1) + (2\times 0) = 5$
$(1\times 0) + (3\times 3) + (2\times -1) = 7$
$(1\times -1) + (3\times 2) + (2\times 1) = 7$

$(0\times 2) + (3\times 1) + (1\times 0) = 3$
$(0\times 0) + (3\times 3) + (1\times -1) = 8$
$(0\times -1) + (3\times 2) + (1\times 1) = 7$

$(2\times 2) + (1\times 1) + (3\times 0) = 5$
$(2\times 0) + (1\times 3) + (3\times -1) = 0$
$(2\times -1) + (1\times 2) + (3\times 1) = 3$

3x2: $\begin{pmatrix} 0 & 2 \\ 1 & -2 \\ 2 & 3 \end{pmatrix} \times \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 6 & 4 \\ 1 & -5 & -5 \\ 9 & 11 & 4 \end{pmatrix}$

$(0\times3)+(2\times1) = 2$
$(0\times1)+(2\times3) = 6$
$(0\times-1)+(2\times2) = 4$

$(1\times3)+(-2\times1) = 1$
$(1\times1)+(-2\times3) = -5$
$(1\times-1)+(-2\times2) = -5$

$(2\times3)+(3\times1) = 9$
$(2\times1)+(3\times3) = 11$
$(2\times-1)+(3\times2) = 4$

- Matrix multiplication IS associative: $(A\times B)\times C \equiv A\times(B\times C)$
- Matrix multiplication IS NOT commutative: $A\times B \neq B\times A$

Matrix x Vector: $\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$

$x' = a_{1,1}x + a_{1,2}y + a_{1,3}z$

$y' = a_{2,1}x + a_{2,2}y + a_{2,3}z$

$z' = a_{3,1}x + a_{3,2}y + a_{3,3}z$

- The complexity of multiplying a $n\times m$ matrix by a $m\times n$ matrix is ab^2 where a is the lower of m and n and b is the larger.

Matrix Transformation

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ = Identity Matrix

$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ = Scales a Vector

$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ = Reflection in x-axis

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ = Reflection in y-axis

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} = \text{Rotation of } \theta^\circ \text{ around } x\text{-axis}$$

$$\begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} = \text{Rotation of } \theta^\circ \text{ around } y\text{-axis}$$

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{Rotation of } \theta^\circ \text{ around } z\text{-axis}$$

$$\begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x + t_x \\ y + t_y \\ z + t_z \\ 1 \end{pmatrix}$$

Translates $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ by t_x, t_y, t_z.

- Transformation can be chained together with multiplication.
    - This is useful for applying the same set of transformations to many vectors, as the combined matrix only has to be calculated once.

Matrix Determinant

2x2:

$$\begin{array}{cc} + & - \end{array}$$
$$\begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix} = 0 + (-1 \times 1) - (2 \times 3) = -7$$

3x3:

$$\begin{array}{ccc} + & - & + \end{array}$$
$$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} = 0 + 1 \times (3 \times 3 - 1 \times 1) - 3 \times (0 \times 3 - 1 \times 2) + 2 \times (0 \times 1 - 3 \times 2) = 2$$

1 × determinant of the pink square
3 × " " " green "
etc...

- The complexity of this grows with $n!$

Matrix Inversion

1. Compute determinant → If zero, it can't be done!
2. Transpose ($A = A^T$)
3. Compute determinant of minor matricies
4. Change signs as required by creating Adj(M)
5. Divide by determinant (from step 1)

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \end{pmatrix}$$

① Determinant $= 0 + 1 \times (3\times1 - 1\times2) - 2\times(2\times1 - 1\times3) + 3\times(2\times2 - 3\times3) = -12$

② $A^T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \end{pmatrix}$ (Reflect along diagonal)

③
$$\begin{pmatrix} 3\times1 - 2\times1 & 2\times1 - 2\times3 & 2\times1 - 3\times3 \\ 2\times1 - 3\times1 & 1\times1 - 3\times3 & 1\times1 - 2\times3 \\ 2\times2 - 3\times3 & 1\times2 - 3\times2 & 1\times3 - 2\times2 \end{pmatrix}$$

$\downarrow$

$$\begin{pmatrix} 1 & -4 & -7 \\ -1 & -8 & -5 \\ -5 & -4 & -1 \end{pmatrix}$$

Minor Matricies $= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \end{pmatrix}$

Determinant of minor matrix for $\square$ is determinant of $\square$.

④ Apply $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 4 & -7 \\ 1 & -8 & 5 \\ -5 & 4 & -1 \end{pmatrix}$

⑤ Divide to give... $\begin{pmatrix} 1 & 4 & -7 \\ 1 & -8 & 5 \\ -5 & 4 & -1 \end{pmatrix} \times \frac{-1}{12}$

Solving Systems of Linear Equations

- How do you solve
$$\begin{aligned} x + 2y - 2z &= -2 \\ -2x + 3y + z &= 2 \\ 3x - y + 3z &= 14 \end{aligned}$$
?

- Normally you'd use substitution and elimination, but that sucks.

- First, translate to Reduced Row-Echelon form:

  - Matrix form: $\left(\begin{array}{ccc|c} 1 & 2 & -2 & -2 \\ -2 & 3 & 1 & 2 \\ 3 & -1 & 3 & 14 \end{array}\right)$

  - We aim for $\left(\begin{array}{ccc|c} 1 & ? & ? & V_x \\ 0 & 1 & ? & V_y \\ 0 & 0 & 1 & V_z \end{array}\right)$

- These operations can be applied:
  - swap position of two rows
  - multiply a row by a non-zero scalar
  - add one row to a scalar multiple of another.

- Pivoting:
  - multiply a row by a non-zero constant and add it to a non-zero multiple of another row
  - use this to replace all non-zero values in the selected column.
  - we can pivot on any value (not just 1)

- Pivot on (0,0):

$$\left(\begin{array}{ccc|c} 1 & 2 & -2 & -2 \\ -2 & 3 & 1 & 2 \\ 3 & -1 & 3 & 14 \end{array}\right) \begin{array}{l} R_1 \\ R_2 - (-2R_1) \\ R_3 - (3R_1) \end{array} \begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \end{array} \left(\begin{array}{ccc|c} 1 & 2 & -2 & -2 \\ 0 & 7 & -3 & -2 \\ 0 & -7 & 9 & 20 \end{array}\right)$$

- Pivot on (1,1):

$$\left(\begin{array}{ccc|c} 1 & 2 & -2 & -2 \\ 0 & 7 & -3 & -2 \\ 0 & -7 & 9 & 20 \end{array}\right) \begin{array}{l} 7R_1 - 2R_2 \\ R_2 \\ R_3 + R_2 \end{array} \begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \end{array} \left(\begin{array}{ccc|c} 7 & 0 & -8 & -10 \\ 0 & 7 & -3 & -2 \\ 0 & 0 & 6 & 18 \end{array}\right)$$

- Pivot on (2,2):

$$\left(\begin{array}{ccc|c} 7 & 0 & -8 & -10 \\ 0 & 7 & -3 & -2 \\ 0 & 0 & 6 & 18 \end{array}\right) \begin{array}{l} 6R_1 + 8R_3 \\ 2R_2 + R_3 \\ R_3 \end{array} \begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \end{array} \left(\begin{array}{ccc|c} 42 & 0 & 0 & 84 \\ 0 & 14 & 0 & 14 \\ 0 & 0 & 6 & 18 \end{array}\right)$$

- Divide:

$$\left(\begin{array}{ccc|c} 42 & 0 & 0 & 84 \\ 0 & 14 & 0 & 14 \\ 0 & 0 & 6 & 18 \end{array}\right) \begin{array}{l} /42 \\ /14 \\ /6 \end{array} \begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \end{array} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$\therefore\ x = 2,\ y = 1,\ z = 3$

Linear Programming

- LP is useful for solving an optimisation problem within certain constraints.
- Very useful in business!

General Form

Maximise $z = W_z \cdot v$

Subject to $W_1 \cdot v \; \{>, <, =\} \; c_1$
$W_2 \cdot v \; \{>, <, =\} \; c_2$
$\vdots$
$W_n \cdot v \; \{>, <, =\} \; c_n$

$\cdot v$ = Dot Product

Example: Maximise $z = 5x_1 + 2x_2$

Subject to $6x_1 + x_2 \leq 6$
$4x_1 + 3x_2 \leq 12$
$x_1 + 2x_2 \leq 4$

Worked Example

Cup of coffee = 7g of beans.
Caffiene content must not exceed 130mg/100ml

Bean A = £6.67 per 1000g
180mg per 100ml from 7g

Bean B = £12.00 per kilo
100mg per 100ml from 7g

What proportion of Beans A and B will minimise cost?

Original Problem

Variables: a = grams of bean A
b = grams of bean B

A_{co} = cost of A per gram = 0.00667
B_{co} = cost of B per gram = 0.012

A_{ca} = caff. per 100ml for A per gram = $180/7$
B_{ca} = caff. per 100ml for B per gram = $100/7$

Objective Function: Minimise $a \times A_{co} + b \times B_{co}$

Constraints: $a + b = 7$ (a cap is 7g of beans)

$a \times A_{ca} + b \times B_{ca} \leqslant 130$ (no more than 130mg/100ml, assuming proportions mix predictably).

Graphical Solution:

① $a + b = 7$

$\Rightarrow a = 7 - b$

② $a \times A_{ca} + b \times B_{ca} \leqslant 130$

$\Rightarrow a \times 25\frac{5}{7} + b \times 14\frac{2}{7} \leqslant 130$

$\Rightarrow 180a + 100b \leqslant 910$

$\Rightarrow 100b = 910 - 180a$

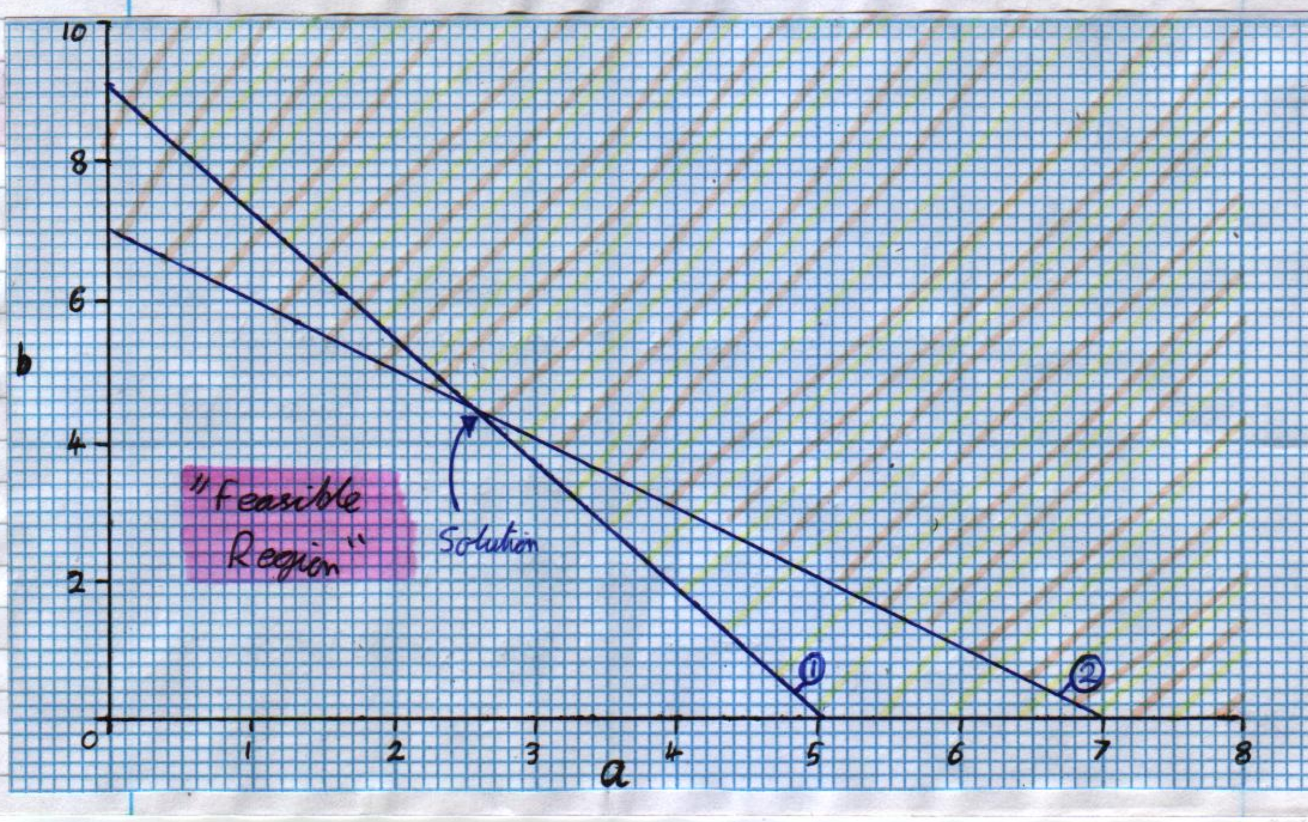
$\Rightarrow b = 9.1 - 1.8a$

① rules out the yellow area

② rules out the orange area.

The solution is on one of the intersections inside the feasible region.

Solution $\Rightarrow$ $a = 2.625g$, $b = 4.375g$



Notes About Solutions

- Solutions always lie at intersections of the lines
  - If there are multiple, check them all to see which is best.
  - They are always on the lines / vertices, because a solution in the "empty space" could be optimised by moving it towards one of the lines.

3 or More Variables

- nD graphs are hard to draw for $n > 2$, so we have to look for a different method...

Integer Programming

- This process allows constraints for many variables to be set out mathematically and then solved by a software solver.

Example: Nurse Rota

Variables: $d_{i,j} \in \{0,1\}$: nurse i is working day shift j

$n_{i,j} \in \{0,1\}$; nurse i is working night shift j

Constraints: (for m days)

$\forall_j \sum_{k=1}^{m} d_{kj} \geq 3, \leq 5$: each day shift has 3-5 nurses

$\forall_j \sum_{k=1}^{m} n_{kj} \geq 2, \leq 4$: each night shift has 2-4 nurses

$\forall i \, \forall_{k \in \{1 \dots m\}} \, n_{i,k} + d_{i,k+1} \leq 1$: no nurse works a night shift and the next day shift

$\forall i \, \forall_{k \in \{1 \dots m\}} \, d_{i,k} + n_{i,k} \leq 1$: no nurse works day and night on the same day.

$\forall i \, \forall_{s \in \{1 \dots m-7\}} \sum_{k=s}^{s+7} n_{i,k} + d_{i,k} \leq 5$: max 5 shifts per nurse per 7 days.

$\sum_{k=1}^{m-1} \forall i \; n_{i,k} \leq 15$: no nurse works more than 15 night shifts across the m days in this rota.

Canonical Form (aka. Standard Form)

- Converting an LP problem to canonical form is the first step of solving a problem with Simplex.
- There are 5 steps:

1. Convert the objective function to maximise
2. Remove constants from the objective problem.
3. Convert $\leq$ constraints to $=$ via slack variables.
4. Convert $\geq$ constraints to $=$ via surplus variables.
5. Make all bounds positive.

Worked Example.

Original:

Minimise: $z = 0.15x_1 + 0.12x_2$

Subject to:

$$0.6x_1 + 0.5x_2 \geq 1.6 \quad (1)$$
$$0.05x_1 + 0.11x_2 \leq 0.3 \quad (2)$$
$$0.18x_1 + 0.05x_2 \geq 0.3 \quad (3)$$
$$x_1, x_2 \geq 0 \quad (4), (5)$$

Step 1.

Maximise $-z = -0.15x_1 - 0.12x^2$ (just $\times(-1)$ if minimise!)

Step 2.

Nothing to do here - z has no constants.

Step 3+4

$$0.6x_1 + 0.5x^2 \geq 1.6 \rightarrow 0.6x_1 + 0.5x_2 - S_1 = 1.6$$
$$0.05x_1 + 0.11x_2 \leq 0.3 \rightarrow 0.05x_1 + 0.11x_2 + S_2 = 0.3$$
$$0.18x_1 + 0.05x_2 \geq 0.3 \rightarrow 0.18x_1 + 0.05x_2 - S_3 = 0.3$$
$$x_1, x_2 \geq 0 \rightarrow x_1, x_2, S_1, S_2, S_3 \geq 0$$

1 - Phase Simplex:

① Convert to standard form (see prev. page) *

② Arrange the initial tableau:
- Put all of the objective function on the LHS
  ie. $z = 120x_1 + 150x_2 \Rightarrow z - 120x_1 - 150x_2 = 0$

③ Select the decision variable
- This is the pivot column
- It is the most negative value in the objective function

④ Perform a min test to select the pivot row
- Divide the ans column by the positive values in the pivot column
- Select the lowest as your pivot row. ← ⑤

⑥ Identify the pivot value at the intersection.

⑦ Add/subtract some multiple of the pivot row from all other non-objective rows to change all values in the pivot column to zero, except for the pivot value.

⑧ Go back to ③ and repeat until...

⑨ All values in the objective row are ≥ 0.
- This is the optimal solution.

⑩ Convert to a solution for variables in the objective function by reading off the rows, ignoring slack/surplus and simplifying.

⑪ Check!!!

* 1-Phase or 2-Phase?

- This is a good point to check if you need to use 1-Phase or 2-Phase.

- Once you add slack/surplus values, check the signs.

- If any are opposite from the RHS, you need 2-Phase

- Eg. $0.7x_1 + 0.15x_2 - s_1 = 0.8 \Rightarrow$ 2Phase!

  $0.2x_1 + 0.41x_2 + s_2 = -1.7 \Rightarrow$ 2-Phase!

(1) Maximise $z = 120x_1 + 150x_2$

Subject to:

$$\left.\begin{aligned} 2x_1 + 3x_2 + s_1 &= 30 \\ \tfrac{2}{3}x_1 + 2x_2 + s_2 &= 16 \\ \tfrac{16}{3}x_1 + 4x_2 + s_3 &= 64 \\ x_1, x_2, s_1, s_2, s_3 &\geq 0 \end{aligned}\right\} \text{Canonical / Standard Form}$$

(6)

(2)

	x_1	x_2	s_1	s_2	s_3	z	ans	min test (4)	
R_1	2	3	1	0	0	0	30	$30 \div 3 = 10$	
R_2	$\frac{2}{3}$	2	0	1	0	0	16	$16 \div 2 = 8$	(5)
R_3	$\frac{16}{3}$	4	0	0	1	0	64	$64 \div 4 = 16$	
R_4	-120	-150	0	0	0	1	0		

(3)

(2) (4)

(6)

(7)	x_1	x_2	s_1	s_2	s_3	z	ans	min test (4)	
$R_1 \to R_1 - \frac{3}{2}R_2$	1	0	1	$-\frac{3}{2}$	0	0	6	$6 \div 1 = 6$	← (5)
$R_2 \to R_2$	$\frac{2}{3}$	2	0	1	0	0	16	$16 \div \frac{2}{3} = 24$	
$R_3 \to R_3 - 2R_2$	4	0	0	-2	1	0	32	$32 \div 4 = 8$	
$R_4 \to R_4 + 75R_2$	-70	0	0	75	0	1	1200		

(3)

(6)

(7)	x_1	x_2	s_1	s_2	s_3	z	ans	min test (4)	
$R_1 \to R_1$	1	0	1	$-\frac{3}{2}$	0	0	6	$6 \div \frac{3}{2} = 4$	
$R_2 \to R_2 - \frac{2}{3}R_1$	0	2	$-\frac{2}{3}$	2	0	0	12	$12 \div 2 = 6$	
$R_3 \to R_3 - 4R_1$	0	0	-4	4	1	0	8	$8 \div 4 = 2$	← (5)
$R_4 \to R_4 + 70R_1$	0	0	70	-30	0	1	1620		

(3)

(7)	x_1	x_2	s_1	s_2	s_3	z	ans	
$R_1 \to R_1 + \frac{3}{8}R_3$	1	0	$-\frac{1}{2}$	0	$\frac{3}{8}$	0	9	
$R_2 \to R_2 - \frac{1}{2}R_3$	0	2	$-\frac{8}{3}$	0	$-\frac{1}{2}$	0	8	
$R_3 \to R_3$	0	0	-4	4	1	0	8	
$R_4 \to R_4 + \frac{15}{2}R_3$	0	0	40	0	$\frac{15}{2}$	1	1680	← (9)

(10) $R_1 : 1x_1 + 0x_2 = 9 \Rightarrow x_1 = 9$

$R_2 : 0x_1 + 2x_2 = 8 \Rightarrow x_2 = 4$

(11) $z = 120 \times 9 + 150 \times 4 = 1680$

2-Phase Simplex

① Add another extra variable to each constraint row.

② We will set everything apart from these extras to zero, give them values, then try to get them down to zero.

③ Create an objective for the extra phase.

④ Re-write your objective in terms of the original problem*.

⑤ Arrange the initial tableau
- We will carry the original objective along for the ride to make things easier later. ← 5a
- We put our entire objective on the LHS as before ← 5b

⑥ Apply simplex as before until the bottom row is all ≥ 0
- Pivot values are shown here circled.

⑦ Finished! If 7a is non-zero then the initial problem is infeasible

⑧ Take the tableau in yellow - this is the initial tableau for the original problem.
- Now apply 1-Phase Simplex.

* In this context, "original problem" means the objective, constraints and first set of slack/surplus variables.

① Maximise $z' = -0.15x_1 - 0.12x_2$ $(z' = -z)$.

Subject to:
$$0.6x_1 + 0.5x_2 - s_1 + s_4 = 1.6$$
$$0.18x_1 + 0.05x_2 - s_2 + s_5 = 0.3$$
$$0.05x_1 + 0.11x_2 + s_3 + s_6 = 0.3$$

$$x_1, x_2, s_1, s_2, s_3, s_4, s_5, s_6 \geq 0$$

③ We want to minimise $s_4 + s_5 + s_6$, so we use "maximise $w = -s_4 - s_5 - s_6$"

④
$$\begin{aligned} w &= -s_4 - s_5 - s_6 \\ &= -(-0.6x_1 - 0.5x_2 + s_1 + 1.6) \\ &\quad - (-0.18x - 0.05x_2 + s_2 + 0.3) \\ &\quad - (-0.05x_1 - 0.11x_2 - s_3 + 0.3) \\ &= 0.83x_1 + 0.66x_2 - s_1 - s_2 + s_3 - 2.2 \end{aligned}$$

⑤ $w - 0.83x_1 - 0.66x_2 + s_1 + s_2 - s_3 = -2.2$ ← 5b

	x_1	x_2	s_1	s_2	s_3	s_4	s_5	s_6	z'	w	ans	min test
R_1	0.6	0.5	-1	0	0	1	0	0	0	0	1.6	$1.6 \div 0 = \infty$
R_2	0.18	0.05	0	-1	0	0	1	0	0	0	0.3	$0.3 \div 0 = \infty$
R_3	0.05	0.11	0	0	1	0	0	1	0	0	0.3	$0.3 \div 1 = 0.3$
5a → R_4	0.15	0.12	0	0	0	0	0	0	1	0	0	
R_5	-0.83	-0.66	1	1	-1	0	0	0	0	1	-2.2	

⑥ →	x_1	x_2	s_1	s_2	s_3	s_4	s_5	s_6	z'	w	ans
$R_1 \to R_1$	0.6	0.5	-1	0	0	1	0	0	0	0	1.6
$R_2 \to R_2$	0.18	0.05	0	-1	0	0	1	0	0	0	0.3
$R_3 \to R_3$	0.05	0.11	0	0	1	0	0	1	0	0	0.3
$R_4 \to R_4$	0.15	0.12	0	0	0	0	0	0	1	0	0
$R_5 \to R_5 + R_3$	-0.78	-0.55	1	1	0	0	0	1	0	1	-1.9

→	x_1	x_2	s_1	s_2	s_3	s_4	s_5	s_6	z'	w	ans
$R_1 \to R_1 - \frac{0.6}{0.18} R_2$	0	0.33	-1	3.33	0	1	-3.33	0	0	0	0.6
$R_2 \to R_2$	0.18	0.05	0	-1	0	0	1	0	0	0	0.3
$R_3 \to R_3 - \frac{0.05}{0.18} R_2$	0	0.096	0	0.278	1	0	-0.278	1	0	0	0.217
$R_4 \to R_4 - \frac{0.15}{0.18} R_2$	0	0.78	0	0.83	0	0	-0.83	0	1	0	-0.25
$R_5 \to R_5 + \frac{0.78}{0.18} R_2$	0	-0.33	0	-3.3	0	0	4.33	1	0	1	-0.6

→	x_1	x_2	s_1	s_2	s_3	s_4	s_5	s_6	z'	w	ans
$R_1 \to R_1$	0	0.33	-1	3.33	0	1	-3.33	0	0	0	0.6
$R_2 \to + \frac{1}{3.33} R_1$	0.18	0.15	-0.3	0	0	0.3	0	0	0	0	0.48
$R_3 \to - \frac{0.278}{3.33} R_1$	0	0.068	0.083	0	1	-0.083	0	1	0	0	0.167
$R_4 \to - \frac{0.83}{3.33} R_1$	0	-0.005	0.25	0	0	-0.25	0	0	1	0	-0.4
⑦ → $R_5 \to - \frac{3.3}{3.33} R_1$	0	0	0	0	0	1	1	1	0	1	0

7a

Data Analysis - Lying w/ Statistics

Why Are Statistics Needed?

- Humans may see patterns where there are none.
- We attribute causality where there is none.
- We over-value confirming evidence.
- We seek out confirming evidence over refuting evidence.
- We suffer confirmation bias.
- We are affected by availability.
- We may seek to blend with social norms.

Sampling

- We need a fair cross-section of the population
- Location may introduce bias
- Collection method may introduce bias
- The questioner may have their own bias
- Sample size is also important: larger is generally better

Confirmation Bias

- Bias is not always deliberate or conscious

- Confirmation bias occurs when we search for, interpret or recall information in a way to favour or confirm our own hypotheses.

    - It also applies to how we respond to presented information: we are less critical of info that confirms our beliefs.

Negative Results

- Publishing negative results is hard, but they can lead to interesting discoveries

- How and why?
    - Conduct 10 surveys, publish only the good results
    - Run an algo 100 times, publish the fastest result
    - People may have different motivations and beliefs.

- One way to combat this issue is the pre-registration of clinical trials.

Averages

- Mean, median and mode could be very different.

- Average is of little value without the range.

- Our thoughs may be influenced by expected values.

- Normal distribution
  - A common continuous probability distribution
  - Often used to represent real-valued random variables whose distributions are not known.
  - 68% of values within 1 standard deviation of the mean
  - 95% " " 2 " " " "
  - 99.7% " " " 3 " " " " "

- Multimodal distributions may appear when the dataset has multiple distinct "zones" or "groups" of data.

- Combining Averages

  - The geometric mean should be used:

    Geometric mean of n numbers $= \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$

  - This avoids the problem of weighting due to sample size.

- Regression towards Mean

  - The more extreme a result is, the more likely it is that it will change on the next attempt, often moving towards the mean.

  - The "better" you are to start with, the harder it is to improve
    - Eg. "value added" in schools

  - Ceiling effects may result in a grouping of data points in the higher categories

Conditional Probability

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$

- Baye's Theorem: $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$ Disease questions

Statistical Significance

- Small sample sises lead to spurious results
- Statistical significance ≠ big results
- Statistical significance = results unlikely to be caused by chance
- Stat. Sig. Test: calc. the possibility that the findings are a result of chance.
  - Typically satisfied if the probability of chance is ≤ 5%
  - The confidence level should be chosen before starting
  - One- or two-tailed?
    - Accept deviations in 1 or 2 directions?
- Results of tests should be combined with care
- All results should be included
- There are multiple test types:
  - Z-test: distribution of the test statistic under the null hypothesis can be approximated by the normal distribution.
  - t-test: if the population variance is unknown and the sample size $n < 30$, this may be appropriate
  - Wilcoxon Ranked-Sign Test: alternative to the paired Student's t-test, a t-test for matched pairs or t-test for dependent samples when the population may not be normally distributed
  - ANOVA: when there are more than two groups to test.

Null Hypothesis

- The default position that there is no relationship between two measured statistics.

- It is not proved, but we can fail to disprove it.

- Sample size may be too small

  - not enough confidence to disprove it ≠ enough confidence to prove it

Correlation vs. Causation

- Spurious correlations can be caused by:
  - Small samples
  - Manipulated scales
  - Data dredging (running too many statistical tests)

- Beware extrapolation

- Correlation ≠ causation

Graphs and Diagrams

- Fox News. Need I say more?

Metrics and Spin

- Choosing certain metrics can skew the impression they give
  - Eg. "average pay" could be a mean skewed by director pay.

- "light and economical" vs. "flimsy and cheap"

- Ratios without absolute values are useless.
  - A 50% increase to a tiny risk is still tiny

- Adjacent sentences can imply incorrect meanings

- Introducing certain metrics can influence the way people act
  - Eg. "value added" in schools - students marked down in earlier years

How to Do An Experiment

- Form a hypothesis
- Carefully design an experiment to test the hypothesis, paying attention to:
  - What may disprove it
  - How the data will be analysed
  - What tests will be used to check the hypothesis validity?
  - What confidence is required?
- Run the experiment and collect data
- Run data analysis
- Report results, even if they do not confirm the hypothesis.

Controls

- Used to monitor placebo effects and/or provide a baseline
- Cannot be chosen retrospectively
  - Those who refuse to participate may share a characteristic
- May be hard to do ethically; eg. "half of these children get a fake vaccine"

Peer Review

- Blind trials, double-blind, etc.
- Important to ensure published literature is reasonably reliable
  - ie. checked by other experts
- Typically 3 reviewers who are experts in the field
  - Beware: they may have their own bias and may not spot all issues
- True test: when the paper is published for the wider academic community to review.
- Don't rely on author conclusions - check the data yourself!
- The key to merit is reproducability.

Recurrence Relations

Basic Sequences

- Arithmetic series: 3, 7, 11, 15, 19, ...
  - Common difference between terms
- Geometric series: 3, 6, 12, 24, 48...
  - Common ratio between terms
- Polynomial series: 5, 10, 17, 26, 37
  - Common second difference is a quadratic series
  - Common third difference is a cubic series, etc

Basic Sequence Formulae

- Arithmetic series: 3, 7, 11, 15, 19
  - nth term: $a_n = a_1 + (n-1)d$
  - Sum of n terms: $a_1 + \dots + a_n = \sum_{k=1}^{n-1} (a_1 + kd) = \frac{n}{2}(2a_1 + (n-1)d)$

  a_1 = first term, d = common difference

- Geometric series: 3, 6, 12, 24, 48
  - nth term: $a_n = a_1 r^{n-1}$
  - Sum of n terms: $a_1 + \dots + a_n = \dfrac{a_1(1-r^n)}{1-r}$

  a_1 = first term, r = common ratio

- Polynomial series: 5, 10, 17, 26, 37
  - Don't worry about these!

Recurrence Relation Basics

- A recurrence relation defines a sequence of values of a multidimensional array of values.
- The definition is recursive (calls itself) and has one or more initial terms (think: base cases).
- Eg: for the fib sequence: $F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$

Example: Power

$$P(x,0) = 1 \quad P(x,1) = x$$
$$P(x,n) = x \times P(x,n-1) \qquad \text{for } n > 1$$

- This is the recurrence relation for raising a number to a positive power

- If we re-arrange it a little, we can determine the complexity:

$P(0) = 1$ — Because a simple return is one operation

$P(1) = 1$

① $P(n) = P(n-1) + 2$ — Because each $P(n)$ stage has 2 arithmetic operations; the × and the -

② $P(n-1) = P(n-2) + 2$ — Because it should hold for any n

③ $P(n) = P(n-2) + 2 + 2$ — Sub $P(n-1)$ from ② into ①

④ $P(n-2) = P(n-3) + 2$ — Same logic as ②

⑤ $P(n) = P(n-3) + 2 + 2 + 2$ — Sub $P(n-2)$ from ④ into ③

⑥ $P(n) = P(n-k) + 2k$ — Generalise from ①, ③ and ⑤

⑦ $P(1) = 1$ — Recall from above

⑧ $P(n) = P(1) + 2k$ — We set $P(n-k) = P(1)$ by taking $n-k=1$. This also gives $k=n-1$

⑨ $P(n) = 1 + 2(n-1)$ — Sub ⑦ and $k=n-1$ into ⑧

⑩ $P(n) = 2n - 1$ — Simplify

- The function has $O(2n-1) = O(n)$ complexity.