A summary of VCE Further Mathematics Units 3 & 4

FURTHER

Casio® ClassPad

UNITS 3 & 4

Notes

✓ Core and all modules
✓ Summary notes
✓ Graduated questions
✓ Comprehensive solutions
✓ Examination advice

Christine McRae
Mark O'Connell

A+

CONTENTS

UNCORRECTED SAMPLE PAGES

CORE:
Data analysis

Univariate data

Types of data

Categorical: A category is recorded when the data is collected. Examples of categorical variables include *gender*, *nationality* and *occupation*.

Numerical: When data is collected a number is recorded. There are two types:

 (i) **Discrete:** The numbers recorded are distinct values, often whole numbers and usually the data comes from counting. Examples include *number of students in a class* or *number of pages in a book*.

(ii) **Continuous:** Any number on a continuous number line is recorded; usually the data is produced by measuring to any desired level of accuracy. Examples include *volume of water consumed* and *height of a growing tree*.

Organising data

Frequency tables: May be used with both categorical and numerical data. Class intervals are used to group continuous numerical data or to group discrete data when there is a large range of values.

Stem-and-leaf plots (or stemplots): Display the distribution of numerical data (both discrete and continuous) as well as the actual data values. Statistics including median, interquartile range and range can be determined from an **ordered** stemplot.

Graphs

Bar chart: Categorical data.
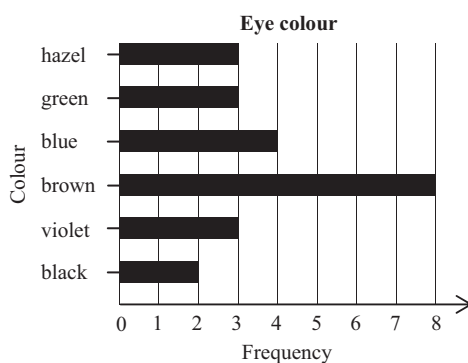
For 'bell-shaped' (normal) distributions:
- 68% of the data will have values within one standard deviation of the mean
- 95% of the data will have values within two standard deviations of the mean
- 99.7% of the data will have values within three standard deviations of the mean.

z scores are used to compare values from different distributions. A raw score, x, from a data set will have a z-score of $\dfrac{x - \bar{x}}{s}$ where $\bar{x}$ is the mean of the data set and s is the standard deviation.

The closer the z-score is to zero, the closer the raw score is to the mean.

Boxplot (or box-and-whisker plot)

A boxplot illustrates the five-number summary [minimum, lower quartile (Q_1), median (Q_2), upper quartile (Q_3), maximum] of a univariate numerical data set and is often used as a visual comparison of similar data sets. The statistics of median, IQR and range can readily be observed from a boxplot.

Outliers are identified using the boundaries $Q_1 - (1.5 \times IQR)$ and $Q_3 + (1.5 \times IQR)$. Any data values outside these boundaries are marked with an asterisk on the boxplot.
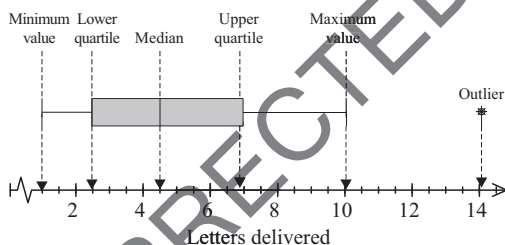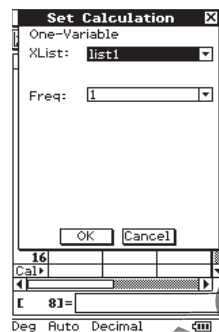

Calculator:
Finding summary statistics and drawing a boxplot.
Choose **Statistics** ⬚ from the **Menu**. Key the data into **list 1**.

Tap **Calc**, **One-Variable**. Complete the screen as shown. Tap **OK**.

Stat Calculation:

Tap **OK**, **SetGraph** and **Setting**. Complete the screen as shown. Tap **Set**.
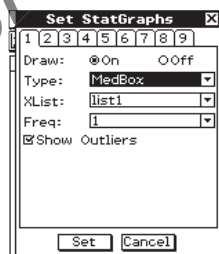
Tap on the graph icon ⬚ to display the boxplot. Tap **Analysis** in the menu bar, select **Trace** then use the cursor key to view the values of the five-number summary displayed on the graph.
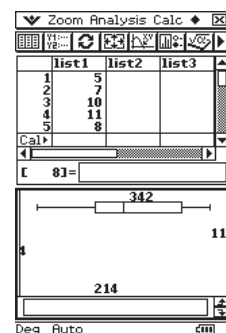
Bivariate data

- **Identify dependent and independent variables:** Answer the question: 'Which of the two variables is likely to depend on the other?'

- **Back-to-back stemplots:** Used to display a relationship between a numerical variable and a two-valued categorical variable. *Example*: Comparison of *weight* (numerical variable) of boys and girls (the alternatives of a two-valued categorical variable, *gender*) in a primary class.
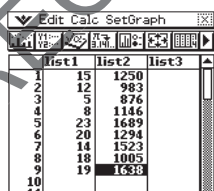- **Parallel boxplots:** Used to display a relationship between a numerical variable and a two or more level categorical variable. *Example*: Comparison of the results of three classes on a statistics test. (Categorical variable is *class* and numerical variable is *result on test*.)
- **Tables** (two-way frequency tables) **and segmented bar charts** (percentage): Used to display the relationship between two categorical variables. *Example*: Comparison of male and female attitudes to Sunday trading. (Categorical variables are *gender* and *attitude*.)
- **Scatterplots:** Used to display the relationship between two numerical variables. *Example*: The relationship between *income* (dependent variable) and *age* (independent variable).
- **Description of scatterplots**
  - Direction: Either positive, negative or random.
  - Form: Either linear or non-linear.
  - Strength: Either weak, moderate or strong.
  - Outliers: Any points that are separated from the general body of points.
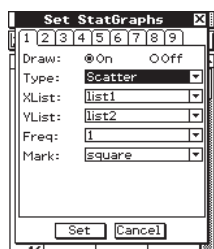
Example on calculator:

The data is the *fat content*, in grams (independent variable), and *energy*, in kilojoules, (dependent variable) of nine foods.

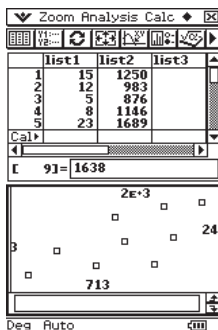To draw a **scatterplot on screen**. Key the data into **list 1** and **list 2**.

Tap **SetGraph** and go into **Setting**. Complete the screen as shown. Tap **Set**.

Tap on the graph icon to display the scatterplot. Tap **Analysis** in the menu bar, select **Trace** then use the cursor key to view the values.
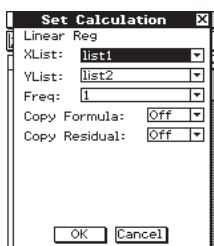
The scatterplot indicates moderate, positive, linear correlation between the variables *energy* and *fat content* with no outliers. As the *fat content* increases the *energy* also increases.

- **Pearson's correlation coefficient** (*r*) is a measure of the strength of the linear relationship between two variables. The value of *r* is between -1 (perfect negative linear relationship) and 1 (perfect positive linear relationship). The key assumption when using Pearson's *r* is that the data is linear.
- **Interpretation** of *r* values. Ignoring the positive or negative sign:
  - 0.75 to 1 is normally regarded as a strong correlation
  - 0.5 to 0.74 is normally regarded as a moderate correlation
  - 0.25 to 0.49 is normally regarded as a weak correlation
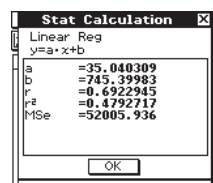 - 0 to 0.24 is normally regarded as very little or no correlation.

Note: Correlation between two variables does **not** imply that a change in one variable causes a change in the other variable.

Example on calculator (to find *r* for the *fat content* and *energy* data):

Tap on **Calc** and choose **Linear Reg**. Complete the **Set Calculation** screen as shown. Tap **OK** to obtain the statistics.

An *r* value of 0.6923 confirms the moderate, positive correlation between the variables as observed from the scatterplot.

- **Coefficient of determination** (r^2)**:** The square of Pearson's correlation coefficient.
  For the *fat content* and *energy* data above, the calculator screen shows an r^2 value of 0.4793. ($0.6923^2 \approx 0.4793$) This statistic shows that 47.93% of the variation in *energy* can be explained by the variation in the *fat content*.

Regression

- **Regression:** Fitting a straight line (often called a linear model) to a scatterplot of bivariate numerical data is called linear regression.

Three methods of regression

  (i) Fitting a line **'by eye'**: Using your judgement for the placement of the line. Used when the graph, but perhaps not the data, is available.
  (ii) The **three-median** regression line is found usually by a graphical technique but the equation can also be found using the calculator if the data is available. It is preferable to use a three-median regression line when there are possible outliers in the data.
  (iii) The **least-squares** regression line is found using the calculator (data must be available) and is a universally accepted method for applying a linear regression line. It should only be used when the graphed data appears to be linear and without outliers.

The regression lines can be used to predict either outside the given data range **(extrapolation)** or within the given data range **(interpolation)**.

Finding the equation of a regression line drawn on a scatterplot

Choose two points on the line (x_1, y_1) and (x_2, y_2).

Use $m = \dfrac{y_2 - y_1}{x_2 - x_1}$ for the gradient.

If the *y*-intercept (*c*) can be read directly from the line (if the *x*-scale starts at zero), then use $y = mx + c$ for the equation.

Otherwise, use the gradient and one point, (x_1, y_1), in $y - y_1 = m(x - x_1)$ for the equation.

Calculator: The equation of the least-squares regression line can be found on the **Linear Reg** screen. Using the previous example of *fat content*

and *energy*, '*a*' is the gradient (slope) and '*b*' is the intercept. The equation in this case is:
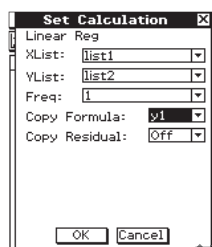*Energy* $= 35.04 \times$ *fat content* $+ 745.4$
or
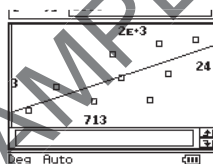*Energy* $= 745.4 + 35.04 \times$ *fat content*
(Statistical equations are often written in this form.)
Tap on **Calc** and choose **Linear Reg**. Change **Copy Formula** to **y1**. Tap **OK** to obtain the statistics. Then tap **OK** again.

Tap on the **graph icon** to display the scatterplot and linear regression line.

- **Interpretation of the gradient or slope** (35.04 in the example)**:** For each increase of one gram of fat, the energy increases by 35.04 kilojoules.
- **Interpretation of intercept:** The energy in a food that contains **zero** grams of fat is 745 kilojoules. If the intercept value is impractical then the regression equation is not suitable for predicting with low values.
- **Residuals:**

$$\text{Residual} = \frac{\text{actual data}}{\text{value}} - \frac{\text{predicted data}}{\text{value}}$$

  A residual plot tests the quality of fit of a model. If the points on a residual plot are randomly scattered then the linear model is appropriate for the data.

Calculator:
Tap on **Calc** and choose **Linear Reg**. Change **Copy Residual** to **list 3**.
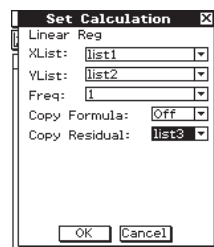
Tap **OK** to obtain the statistics. Then tap **OK** again. The residuals fill **list 3**.

|	list1	list2	list3
1	15	1250	-21
2	12	983	-182.8
3	5	876	-44.6
4	8	1146	120.27
5	23	1689	137.67
6	20	1294	-152.2
7	14	1523	287.03
8	18	1005	-371.1
9	19	1638	226.83
10			

To plot the residuals:
Tap **SetGraph** and go into **Settings**. Turn **Off** graph 1. Tap on graph 2 and complete the settings as shown.
Tap **Set**.

Set StatGraphs
1 2 3 4 5 6 7 8 9
Draw: ◉On ○Off
Type: Scatter
XList: list1
YList: list3
Freq: 1
Mark: square

[Set] [Cancel]

Tap on the **graph icon** to obtain the residuals plot.
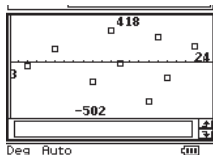
The points are randomly scattered around the zero line, so the linear model is a good fit. If the residual plot shows a systematic pattern then this suggests that a non-linear model is more appropriate. A transformation can be made to one of the axes scales.

• **Estimation of the equation of the least-squares regression line:** The equation is of the form $y = mx + c$ where $m = \dfrac{rs_y}{s_x}$ and $c = \bar{y} - m\bar{x}$.
  $\bar{x}$ and $\bar{y}$ are the means of x and y respectively.
  s_x and s_y are the standard deviations of x and y respectively.
  r is Pearson's correlation coefficient.

• **Transformations:** When a linear model is not appropriate then a set of data can be transformed to linear data by applying a transformation to one of the axes.

The transformation(s) to make the data linear will depend on the underlying shape of the points on the original scatterplot:
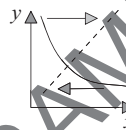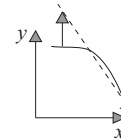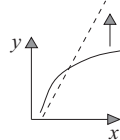
• x^2 stretches the x-scale
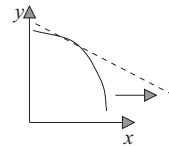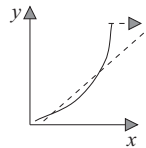
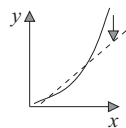• log x compresses the x-scale
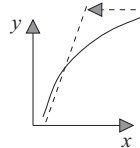
• $\dfrac{1}{x}$ compresses the large x values and stretches the small x values
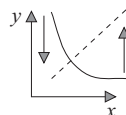

• y^2 stretches the y-scale

• log y compresses the y-scale

• $\dfrac{1}{y}$ compresses the large y values and stretches the small y values.
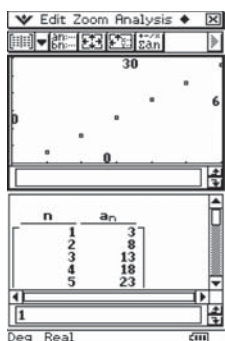

Time series

If time is the independent variable and data is collected over successive intervals of time then the data is called **time series** data. A line graph

Graphs of first-order difference equations
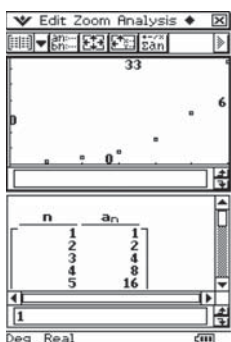(Use **Sequence** ⌨ on the calculator **Menu**.)

Arithmetic sequences:
Points on the graph are
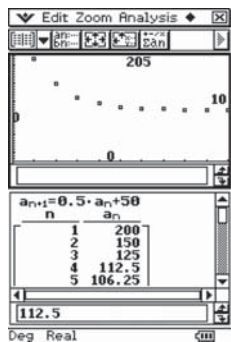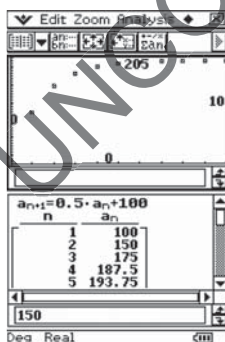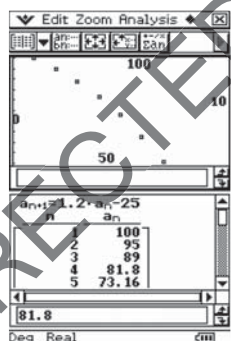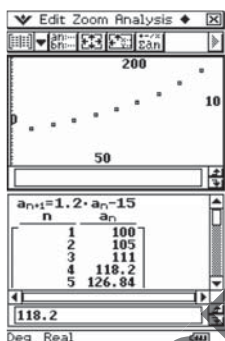in a straight line.

Geometric sequences:
Points on the graph are
not in a straight line.

Sequences that are neither arithmetic nor
geometric (examples):

Points on the graphs are not in a straight line.

The Fibonacci sequence:

Fibonacci's sequence is generated from the
difference equation $t_{n+2} = t_{n+1} + t_n$, where $t_1 = t_2 = 1$.
This should be read as: 'Each new term from
the third term onwards is the sum of the two
previous terms. The first two terms are both one.'
So Fibonacci's sequence is: 1, 1, 2, 3, 5, 8, 13, 21 . . .
Fibonacci's sequence is an example of a second-
order linear difference equation. These equations
have the form $t_{n+2} + at_{n+1} + bt_n = c$, where a, b and c
are constants. To find the terms of a second-order
difference equation, two starting values are
needed.
Second-order linear difference equations can
generate sequences that are arithmetic, geometric
or neither.

Examples:

(i) $t_{n+2} = 6t_{n+1} - 9t_n$, where $t_1 = 1$ and
    $t_2 = 3$ will generate the sequence 1, 3, 9, 27,
    81, . . . which is a *geometric sequence* with first
    term 1 and common ratio of 3.

(ii) $t_{n+2} = 2t_{n+1} - t_n$, where $t_1 = 2$ and $t_2 = 4$
    will generate the sequence 2, 4, 6, 8, 10,
    12 . . . which is an *arithmetic sequence* with
    first term 2 and common difference of 2.

(iii) $t_{n+2} = 3t_{n+1} + 2t_n$, where $t_1 = 1$ and $t_2 = 2$, will
    generate the sequence 1, 2, 8, 28, 100, . . .
    which is neither arithmetic nor geometric.

On the calculator:
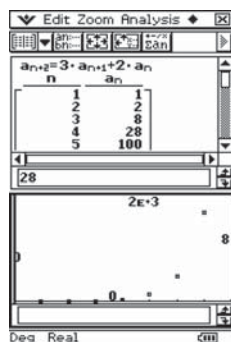
Module 1: Number patterns

Multiple-choice questions: 55 questions

Arithmetic and geometric sequences:
35 questions.
Solutions to this section start on page 202.

QUESTION 91✪

An arithmetic sequence has a first term of 32
and a common difference of $\frac{1}{2}$. The sum of the
first five terms of this sequence is:

A 22
B 62
C 64
D 155
E 165

QUESTION 92✪

Which of the following is a geometric sequence
with $r > 0$?

A -2, 4, -8, 16, . . .
B $\frac{1}{2}, \frac{1}{6}, \frac{1}{8}, \frac{1}{16}, \cdots$
C 1000, 100, 10, 1, . . .
D 14, 11, 8, 5, . . .
E 2, 6, 24, 120, . . .

QUESTION 93✪

Which of the following is **not** a term in the
geometric sequence: 1, $\frac{1}{2}, \frac{1}{4}, \ldots$?

A $\frac{1}{6}$
B $\frac{1}{8}$
C $\frac{1}{16}$
D $\frac{1}{64}$
E $\frac{1}{256}$

QUESTION 94✪

An arithmetic sequence is defined by $t_n = 2n + 1$.
Values for the first term and the common
difference, respectively, of the sequence are:

A $a = 2$ and $d = 1$
B $a = 3$ and $d = 1$
C $a = 3$ and $d = 2$
D $a = 1$ and $d = 2$
E $a = 2$ and $d = 3$

QUESTION 95✪

Two different sequences are shown in list 1 and
list 2 of the calculator screen dump below. Each is to
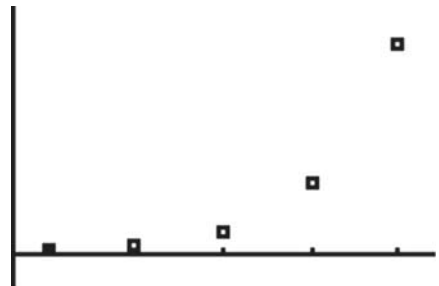be categorised as arithmetic, geometric or neither.

These two sequences are respectively:

A arithmetic and arithmetic
B arithmetic and geometric
C arithmetic and neither
D geometric and neither
E neither and neither

QUESTION 96✪

The first five terms of a sequence are plotted on
the graph below.

This sequence could be:

A an arithmetic sequence with $d = -2$
B an arithmetic sequence with $d = 2$
C a geometric sequence with $r = -2$
D a geometric sequence with $r = \frac{1}{2}$
E a geometric sequence with $r = 2$

QUESTION 97✪✪

A sequence has its first three terms as: a, 4, b.
If this sequence is arithmetic, values of a and b,
respectively, for this sequence could be:

A 2 and 8
B 3 and 7
C 1 and 16
D -2 and 10
E -1 and -4

Example:

Malcolm invests $5000 at 7.5% per annum compounding monthly. Each month, immediately after the interest is calculated, he deposits $400. Determine the value of the investment after 2 years.

$P = 5000$ $\qquad R = 1 + \dfrac{\frac{7.5}{12}}{100} = 1.006\,25$

$n = 2 \times 12 = 24$ $\qquad Q = 400$

$A = 5000 \times 1.006\,25^{24} + \dfrac{400\left(1.006\,25^{24} - 1\right)}{1.006\,25 - 1}$

$A = \$16\,129.15$

On the calculator, the sign of the present value and the payment must both be the same. In this case, the original investment and the payments are considered to be 'outgoings' and so have negative signs. (See page 115 for an explanation of symbols.)

Tap on **Financial** on the **Menu**. Tap on **Calculations** and choose **Compound Interest**. Complete the screen as shown and then tap on the **FV** box to find the value of the investment after 2 years.
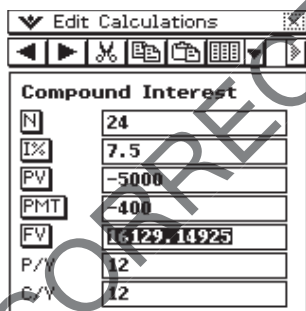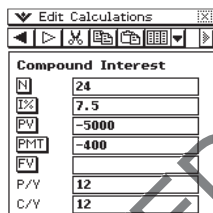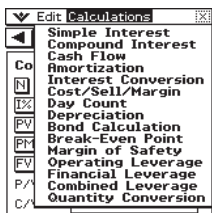
Perpetuities

A perpetuity is an investment that provides a regular payment (interest) on an indefinite basis without affecting the original value of the investment. The perpetuity is the interest earned by the principal in the specified time period. The most likely use of a perpetuity is as a pension which a retired person will receive for the remainder of his or her life.

Formula:

$P = \dfrac{100Q}{r}$

P: The initial principal invested

Q: Perpetuity paid per time period

r : Percentage interest rate per time period

Example:

Albert invests $500 000 at 4.9% per annum. Determine the fortnightly perpetuity he will receive.

$P = 500\,000 \qquad r = \dfrac{4.9}{26} = 0.188\,462$

$P = \dfrac{100Q}{r} \Rightarrow Q = \dfrac{Pr}{100}$

$\qquad = \dfrac{500\,000 \times 0.188\,462}{100}$

$\qquad = \$942.31$

Reducing balance loans and annuities

Characteristics of a reducing balance loan:
- regular payments are made throughout the term of the loan
- the interest is calculated for each time period on the reduced balance owing
- with each repayment, the proportion of interest and reduction to principle varies, with more interest being paid early in the loan.

Formula:

$A = PR^n - \dfrac{Q\left(R^n - 1\right)}{R - 1}$, where $R = 1 + \dfrac{r}{100}$

A: The amount owing

P: The initial amount borrowed

R: The multiplying factor

r : Percentage interest rate per time period

n : Number of time periods

Q: Amount of the regular repayment

If the repayment figure, *Q*, is unknown, the formula can be transposed:

$Q = \dfrac{\left(PR^n - A\right)\left(R - 1\right)}{\left(R^n - 1\right)}$

Use of the TVM solver on the calculator

While an understanding of the formula is very helpful, on most occasions answers can be determined directly from the calculator. Explanation of the various terms on the following screen is best done by way of example.

Tap on the 2×2 matrix symbol and a 2×2 matrix will appear on the screen.
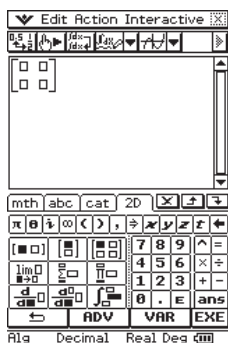
If you want a 3×2 matrix, tap once on and this will produce another row.
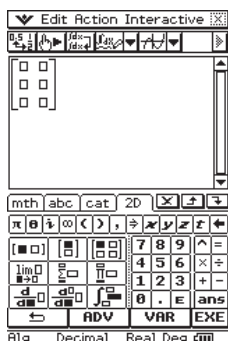
Tap on to produce another column. Tap on the first element in the matrix and tap the number on the keyboard, or use the number keys, to enter the first element shown on the screen below. Use the cursor key arrows to move to the next element and continue entering the elements to match the matrix on the screen below.

To save the matrix, tap outside the matrix and **save** the matrix as a variable (tap *VAR*), *A* in this case.
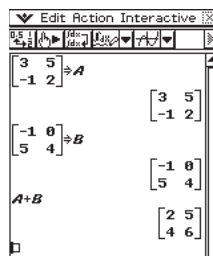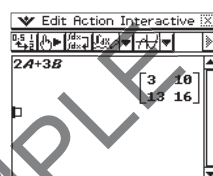
Tap [EXE] and the matrix is saved as *A*.
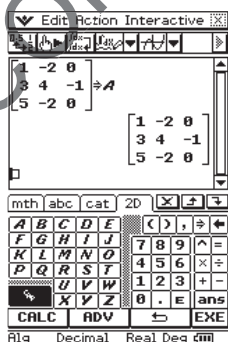
Operations on the calculator:

Once you enter and save your new matrices **A** and **B**, make sure you use the variables *A* and *B* in each calculation and then tap [EXE].

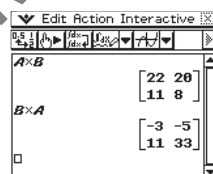(i) Addition.

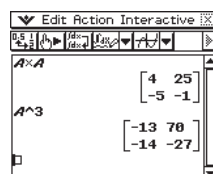(ii) Multiplication by a scalar and calculation of expressions.

(iii) Multiplication.

Powers of a matrix

Powers of a **square** matrix: $\mathbf{A}^2 = \mathbf{A} \times \mathbf{A}$;
$\mathbf{A}^3 = \mathbf{A} \times \mathbf{A} \times \mathbf{A} = A\wedge 3$ on the calculator, etc. The calculator can be used to find powers of a matrix. Make sure you are using the variable *A*.

Inverse matrices

If **A** is a square matrix, its **inverse matrix** is denoted by $\mathbf{A}^{-1}$.

$\mathbf{A}^{-1}$ is the matrix such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, where **I** is the identity matrix. This property of inverse matrices makes it possible to solve matrix equations.