**Part I Mathematical Background**

**Linear Algebra**  
**Def: A vector is a collection of numbers in R. (or C, F)**Ex. , ,   
 column vector row vector  
Throughout the semester, "vector" = "column vector", "vectorT" = "row vector",

Ex.

**Def: Two operations:  
1.Vector addition:**   
**2.Scalar multiplication: For** , , **&**   
Combining 1 & 2, we have a linear combination of vectors of vectors

**Def: The inner product of two vectors  & is**

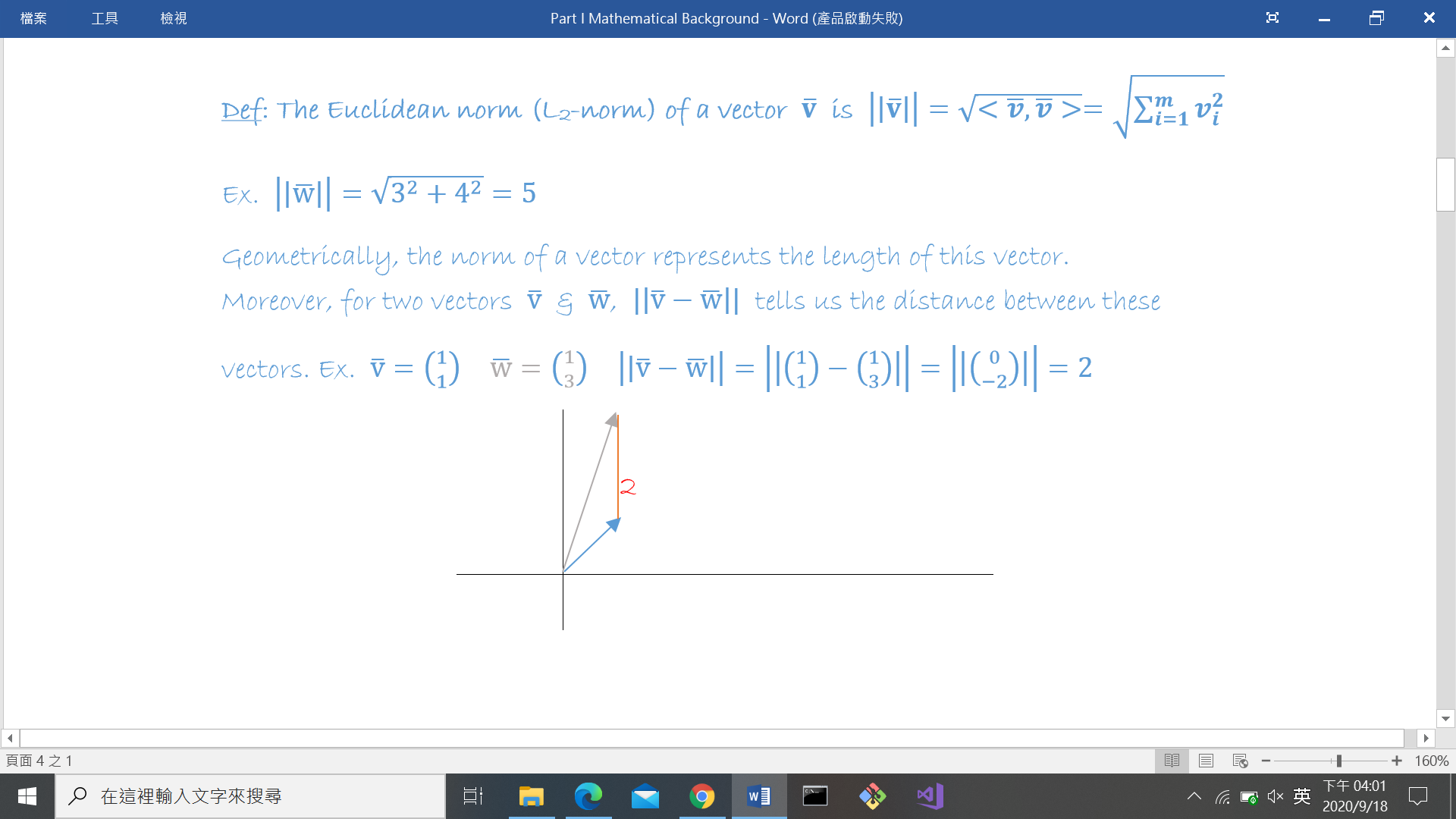
Ex.

**Def: The Euclidean norm (L2-norm) of a vector is**

Ex.

Geometrically, the norm of a vector represents the length of this vector.

Moreover, for two vectors & , tells us the distance between these vectors. Ex.



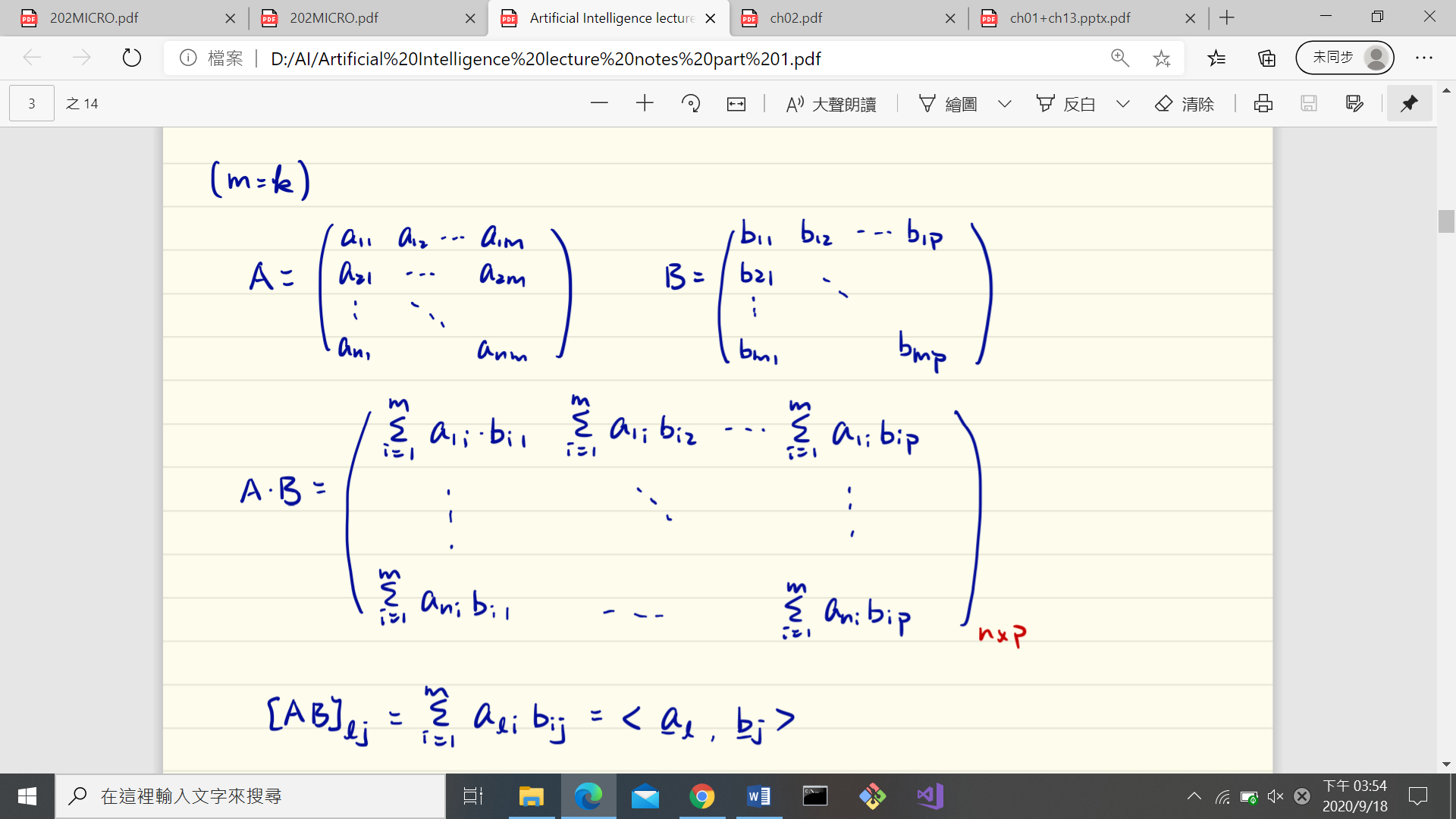
**Def: An matrix is rectangular array of values arranged into n rows and m columns.**

Ex.

, are matrices, is a matrix.

Two matrices with the **same dimension** can be added elementwise to form a new matrix.

A (nm) and B (kp) can be multiplied into **AB** if and only if **m = k**, and can be multiplied into **BA** if and only if **p = n**.



where is the ith row of and is the kth column of .

Ex.

Note that in general. Multiplication of matrices does **NOT** satisfy **commutative** law. They even may not have the same dimensions.

**Def: The transpose of a matrix A is the matrix AT where the columns of AT are the rows of A.**

Ex.

**Def: The identity matrix is a square matrix () whose**

**Def: A square matrix is said to be non-singular (invertible) if there exists a matrix such that . Otherwise, it is singular.**

Finding A-1 (if exists) can be done by **Gauss-Jordan elimination**, which involves complexity of **O(n3).**

Application: , solve for .

**Probability**

**Def: A random variable (RV) is a function**

**Ω: Sample space consisting of all the possible outcomes.**

Ex. Toss a coin, Ω = {Head, Tail}

Win $1 if the outcome is H

Lose $1 if the outcome is T

**Def: An admissible subset of Ω is an event.**

Ex: Roll a die. Ω = {1, 2, 3, 4, 5, 6}

Event Big = {4, 5, 6} Small = {1, 2, 3}

**Def: For a countable Ω, a probability distribution called probability mass function (PMF) specifies the likelihood of any event. It assigns any a real number such that**

**1) for every**

**2) (normalization)**

**3) For and ,**

**Def: Let be an event such that . A conditional probability distribution assigns to every event a number**

Ex. The die is fair. P(i) = 1/ 6 I {1, 2, 3, 4, 5, 6}

P(Big) = 1/2

P(4|Big) = (1/6) / (1/2) = 1/3

P(1|Big) = 0

**For uncountably infinite Ω (for example [a, b]) the probability of any particular point is 0. PMF is not well-defined.**

**Def: For a continuous RV X, the cumulative distribution function (CDF) is , which records the probability of X smaller than x.**

**If F(x) is differentiable, we can define the probability density function (PDF) as or**

**Def: For a RV X, the expectation is defined as (1st moment), or (2nd moment), capturing the ensemble average of X.**

**The variance of X is defined as measuring the spread of X out from .**

**Some frequently encountered RVs**

**Bernoulli: X~Bern(P)**