MAT133 Finite Mathematics

Chapter 2: Systems of Equations: m equations in n variables

Graphing, Substitution or Elimination (manual & messy)

Matrix (Augmented **m by n+1 matrix** – see below): **RREF**
Row Operations: Multiply by #, Swap, Add.
Well-determined (m=n): usually **unique solution**; with RREF
Diagonal of 1's on left with solution for each variable on right.
Under-determined (m<n): usually **infinity of solutions**, with RREF
Multiple #'s in some rows for dependency relations.
Over-determined (m>n): usually **no solutions**, with RREF
Row (last) of all 0's on left with non-zero on right

Inverse Matrix(see below): Solve $\mathbf{M \bullet X = Y}$, by $\mathbf{X = M^{-1} \bullet Y}$.

Chapter 3: Matrix Algebra

Matrix is: $\mathbf{M = (m_{ij})}$; **i = 1 to m** rows, **j = 1 to n** columns
Shape of matrix is **"m by n"** or "**m**×**n**".
"m by 1" is a **row vector**;   "1 by n" is a **column vector**.
Swapping rows and columns yields the **transpose** matrix: $\mathbf{M^T}$.
The transpose of a row vector is a column vector and vice versa.

Matrix Addition (of same shape) and **"scalar"** multiplication (any shape)
Combine or convert data tables of "like data"

Matrix Multiplication: **"m by p" • "p by n" = "m by n"**
(sum of products across rows (left) and down columns (right))
(i) **Allocates** quantities by proportions (in the matrix)
(ii) Matrix defines a function:
Matrix on left defines a function on row vectors:
"m by n" • "n by 1" yields "m by 1"
Matrix on right defines a function on column vectors:
"1 by m" • "m by n" yields "1 by n"
Matrix multiplication is composition of functions
Square matrices may be invertible with inverses.

$$\mathbf{M \bullet M^{-1} = Id = M^{-1} \bullet M}$$ **, Id is Identity (1's on diagonal)**

Laws of Matrix Algebra: Matrices may be manipulated with these three operations, obeying **laws of algebra** similar to the usual ones.

<u>TI-84 Matrix Capabilities:</u>

2^{nd}-Matrix : choose and/or display
2^{nd}-Matrix-Edit : enter data. May often need multiple matrices.
2^{nd}-Matrix-Math-rref : put in RREF and solve the related system.
2^{nd}-Matrix-1, 2^{nd}-Matrix-Math-T : compute transpose, $\mathbf{A}^T$
#**, *, 2^{nd}-Matrix-1** : compute scalar product #**B**
2^{nd}-Matrix-1, +, 2^{nd}-Matrix-2 : compute matrix sum **A+B**
2^{nd}-Matrix-1, *, 2^{nd}-Matrix-2 : compute matrix product **A•B**
2^{nd}-Matrix-1, x^{-1} : compute inverse matrix, $\mathbf{A}^{-1}$

<u>Game Theory:</u> Zero sum games between 2 players A & B

Payoff m by n Matrix: A's payoff for each pair of **A's m moves** and **B's n moves**, with entries = A's winnings (>0) or losses (<0) for that pair.

Random strategy for A: row vector with the frequency of each move.
Random strategy for B: column vector with the frequency of each move.

Expected Value of 2 Strategies: A's average winnings over many games.
EV = (A-Strategy)•(Payoff)•(B-Strategy) = ± number

Maximum Pure A-Strategy against a **known B strategy**, compute:
(Payoff)•(B-Strategy) = m by 1 vector,
find row with largest entry, A plays that move 100% of the time.

Maximum Pure B-Strategy against a **known A strategy**, compute:
(A-Strategy)•(Payoff) = 1 by n vector,
find column with least (most negative) entry, B plays that move 100%.

Optimal, Minimax Strategy: The random strategy which minimizes losses against all the opponents possible strategies.

For 2 by 2 games:
A) compute: **(x 1-x)•(Payoff) = (f(x) g(x))**, choose x by f(x)=g(x)
B) compute: **(Payoff)•$(y\ \ 1\text{-}y)^T$ = $(f(y)\ \ g(y))^T$**, choose y by f(y)=g(y)

For more complicated games, **Linear Programming** is usually required to compute the Optimal Minimax Strategy; although, **Dominance Simplifications** or **Saddle Points (row minimum & column maximum)** will sometimes allow optimal minimax strategies to be found.

Chapter 4: Linear Programming – An Optimization Method

To optimize a linear (objective) function in n-variables: $\mathbf{x_1, x_2, \ldots, x_n}$;

$$P(x_1, x_2, \ldots, x_n) = (p_1, p_2, \ldots, p_n) \bullet (x_1, x_2, \ldots, x_n)^T$$

Subject to constraints defined by inequalities:

$$A: (A_1, A_2, \ldots, A_n) \bullet (x_1, x_2, \ldots, x_n)^T \leq A$$
$$B: (B_1, B_2, \ldots, B_n) \bullet (x_1, x_2, \ldots, x_n)^T \leq B$$
$$C: (C_1, C_2, \ldots, C_n) \bullet (x_1, x_2, \ldots, x_n)^T \leq C$$

…..

Which define a bounded region, the maximum and minimum values will occur at the "corner" points of the region.

Chapter 5: Mathematics of Interest – The Cost of Renting Money

A **present value** of money:	**PV**	
is invested/borrowed at:	**r%**	annual interest rate
the **interest** after 1 year is:	$PV \bullet r$	
The **future** value after t years:	$PV \bullet (1 + rt)$	**simple interest**
	$PV \bullet (1 + r)^t$	**compound interest**
Compounded m times per year: (rate of r/m applied m^th^ly)	$PV \bullet (1 + r/m)^{tm}$	(monthly is m=12)
Compounded continuously:	$PV \bullet e^{tr}$	$e = \lim (1 + 1/m)^m$

Loan:
PV = L = Loan amount
P = payment (usually monthly)
r = rate per payment period (usually annual rate / 12)
n = number of payments (end of each period)

$$P = \frac{L \bullet r}{1 - (1 + r)^{-n}}$$

Excel Basics:

Cell addresses are a column letter and row number: **"L#"**
Fixed Cell addresses: **"L$#"**
Calculations begin with: **"="**
Calculations use numbers, cell addresses, **+, -, *, /, ^**
To sum a column(=L) range: **"=Sum(Ln:Lm)"**
To sum a row(=n) range: **"=Sum(An:Zn)"**

Chapter 6: Set Theory and Counting (Cardinality)

A **set** is a collection of **elements.** An element is a member of a set.

x ε S means **x** is an element of **S**.

Sets are lists {…} or Truth Sets {x | P(x)} or one of these:

Naturals N $\quad$ **N_n = { 1, 2, 3, …, n }**

Integers Z $\quad$ **Z_n = { 0, 1, 2, …, n – 1 }**

Rationals Q $\quad$ **Reals R** $\quad$ **Complex #s C**

$N^+ = Z^+$, Q^+, R^+ just the positive numbers in each.

Given two sets, **A** and **B** :

Equal: $\quad$ **A = B ≡ (x ε A ↔ x ε B)**

Subset: $\quad$ **A ⊆ B ≡ (x ε A → x ε B)**

Proper Subset: **A ⊂ B ≡ (A ⊆ B ∧ ¬A = B)**

The Null Set = ∅ = { } **has no elements. For all sets** S, ∅ ⊆ S .

For a set, **S,** the set of all subsets of **S** is the **Power Set of S:**

$P(S) = 2^S$; $\quad$ **∀S (∅ ε P(S)).**

Definitions: Given two sets, A and B:

Product (ordered pairs): $\quad$ **A × B = { (a,b) | a ε A ∧ b ε B }**

$\quad$ **A^2 = A × A = { (a,b) | a,b ε A }**

Union: $\quad$ **A ∪ B = { x | x ε A ∨ x ε B }**

Intersection: $\quad$ **A ∩ B = { x | x ε A ∧ x ε B }**

Difference: $\quad$ **A − B = { a | a ε A ∧ ¬a ε B }**

Subsets of **universal set, U**: $\quad$ **$A^C = \overline{A}$ = U – A , complement of A.**

DeMorgan: $\quad$ $\overline{(A \cap B)} = (\overline{A} \cup \overline{B})$ $\quad$ $\overline{(A \cup B)} = (\overline{A} \cap \overline{B})$

Associative and Distributive Laws (prove with Venn Diagrams):

(A∪B)∪C = A∪(B∪C) $\quad$ **(A∩B)∩C = A∩(B∩C)**

A∪(B∩C) = (A∪B)∩(A∪C) $\quad$ **A∩(B∪C) = (A∩B)∪(A∩C)**

A **function** from a set **X** to a set **Y**, **f:X —> Y**, is a rule which assigns to every x precisely one, **y=f(x)** . **Graph of y=f(x) is:**

G(f) = { (x,y) ε X×Y | y = f(x) } ⊆ X×Y

For two functions, **f,g:X—>Y**, with **Y** a set of numbers, define:

f±g(x) = f(x)±g(x) **f·g(x) = f(x)·g(x)** **f/g(x) = f(x)/g(x)**

For two functions, **f:X—>Y** and **g:Y—>Z,** define the **composition:**

g°f: X—>Z as g°f(x) = g(f(x))

Two functions, **f:X—>Y** and **g =** $\mathbf{f^{-1}}$**:Y—>X** , are **inverses** iff

$\forall$**x ε X (g°f(x) = x)** $\wedge$ $\forall$**y ε Y (f°g(y) = y).**

"How Many?" - Cardinality or Counting

If **A** is a set, **#(A) = n(A) =** number of elements in **A** .

Principles of Counting: Conceive of counting as a process, and

Sum Rule: multi-case process: total = sum of cases
Product Rule: multi-step process: total = product of steps
Re-order: Count in specific order times number of ways to reorder.

Multiple Sets (not necessarily disjoint cases):

#(AUB) = #(A) + #(B) − #(A∩B)
#(AUBUC) = #(A) + #(B) + #(C)
− #(A∩B) − #(A∩C) − #(B∩C) + #(A∩B∩C)
#(A−B) = #(A) − #(A∩B) $\mathbf{\#(A^C) = \#(U) - \#(A)}$

N Elements – Select K of Them

A **permutation** is selecting by **picking** in a specific order. Permutations are shown as lists enclosed within <...>.
A **combination** is selected by **choosing** without an order, a subset. Combinations are shown as lists enclosed within {...}.

If set A has **N** elements:

the number of permutations with **N** elements is:

N factorial = N! = N*(N-1)*...*2*1 ; 0! = 1

the number of permutations with **K** selections is:

$\mathbf{P(N,K) = {}_NP_K = N! / (N - K)!}$

the number of combinations with **K** selections is:

$\mathbf{C(N,K) = {}_NC_K = N! / (K! (N - K)!) = N}$

Chapters 7&8: Probability

Begin with a **random process** of **outcomes** in a **sample space**, **S** . Each execution of the random process is called a **trial**.

An **event, A,** is a result of the random process. A is a subset of S, $A \subset S$.
An **outcome** is an atomic event. An outcome is an element of S.

Probability measures the **frequency** or **likelihood** of an event.
If A is an event, the probability of A is **P(A)** with $\mathbf{0 \le P(A) \le 1}$:

P(certain event) = P(something happens) = **P(S) = 1**
P(impossible event) = P(nothing happens) = **P(Ø) = 0**
P(rare event) < 0.05 (sometimes < 0.025 or < 0.01)

If all outcomes of the process are equally-likely, then:

P(A) = (# outcomes in A) / (total # of outcomes).

For any process, for a large number of trials, then approximately:

P(A) = (# trials producing A) / (total # of trials).

The **Law of Large Numbers** says the two definitions are consistent.

Given any two events, A, B:

AorB = $A \cup B$ = all outcomes in A or B or both
A&B = $A \cap B$ = all outcomes in both A and B
A^C = “Not A” = Complement of A = all outcomes not in A.

The **laws of probability** are rules for calculating unknown probabilities from known probabilities. The fundamental law is:

$$\mathbf{P(A \text{ or } B) = P(A) + P(B) - P(A\&B)}$$

It follows that: $P(A) + P(A^C) = 1$

Two events, A, B, are **disjoint** if: **A&B = Ø or P(A&B) = 0**
(Think of disjoint events as “mutually exclusive”.)
Two events, A, B, are **independent** if: **P(A&B) = P(A)*P(B)**
The **conditional probability** of “B given A”: **P(B|A) = P(A&B) / P(A)**

If A,B are independent, then: P(B|A) = P(B) , thus, independent events do not affect each other.

Discrete Random Variable is the result of a random process with discrete numeric outcomes, X, where each outcome has probability:

$P(x)$ = $P\{X=x\}$ = $P(X=x)$ satisfying: **$P(x) \geq 0$** and **$\Sigma P(x) = 1$** .

P is the probability distribution function. The probability of an event is the sum of the probabilities of the outcomes which comprise it.

The **mean** or **average** or **expected value** of a discrete random variable:

$$\overline{x} = \mu = \mu_x = E(X) = \Sigma\, x\, P(x)$$

The expected value is the expected average of a large number of trials. It may be used to determine if a bet or game is fair: $E(X) = 0$; or if a possible gain is worth the risk of loss: $E(X) > 0$ or < 0.

The **variance** **(σ^2)** and **standard deviation** (σ) is:

$$\sigma^2 = \sigma_x^2 = \Sigma\, (x - \mu_x)^2\, P(x) \quad \text{and} \quad \sigma = \sqrt{\sigma^2}$$

The standard deviation measures average closeness to the mean. It is used to estimate how likely a random trial will be close to the mean.

To calculate μ's and σ's on the TI-83, enter the probablity distribution like a frequency distribution and use:

STAT - CALC - 1: 1-Var Stats (L_1,L_2)

Binomial probability distribution is random variable, X, based on:

1. **n** independent trials with binary outcomes, **{Success, Failure}**;
2. for each trial, **P(Success) = p** and **P(Failure) = q = 1 – p;**
3. **X** counts the number of successes:

$$P(X=x) = P(x) = {}_nC_x\, p^x\, q^{(n-x)} = \left({}^{n!}/_{x!\,(n-x)!} \right) p^x\, q^{(n-x)}$$

For a binomial random variable: $\mu = np$ $\sigma^2 = np(1 - p)$

On the TI-83, calculate $P(X = x)$ as: 2nd-Vars, binompdf(n, p, x)

On the TI-83, calculate $P(X \leq x)$ as: 2nd-Vars, binomcdf(n, p, x)

Continous Random Variable is the result of a random process with continuous numeric outcomes, **X** , in a domain **D** . It has a probability density function, **p(x)** , defined on D and satisfying:

$p(x) \geq 0$ **Area(under p(x) over all of D) = 1**

The probability that $a \leq X \leq b$ = $P(a \leq X \leq b)$ = $P\{ a \leq X \leq b \}$ = Area under p(x) from a to b. For continuous random variables, **$P(X=a) = 0$.**

Normal probability density (bell curve - most data) is defined on {all x} as:

$$p(x) = \left({}^{1}/_{\sigma\sqrt{2\pi}} \right) e^{-(x-\mu)^2 / 2\sigma^2}$$

The standard normal distribution (**Z-score** table) has $\mu = 0$ and $\sigma = 1$.
For arbitrary normal random variable, translate x into a Z-score:
$Z(x) = (x - \mu) / \sigma$ and lookup area in the Z-score table.

Prob{ X < a } = area for Z(a). Prob{ X > a } = 1 – area for Z(a).
Prob{ a < X < b } = area for Z(b) – area for Z(a).

Remember Rules of Thumb

Prob $\{\mu-\sigma < X < \mu+\sigma\} \approx 0.68$
Prob $\{\mu-2\sigma < X < \mu+2\sigma\} \approx 0.95$
Prob $\{\mu-3\sigma < X < \mu+3\sigma\} \approx 0.995$

On the TI-83, calculate the Prob{ a < X < b }:
2nd-Vars, normalcdf(a, b, μ, σ);
for negative infinity, {X < b}, use at least 5σ below μ;
for positive infinity, {X > a}, use at least 5σ above μ.

To find the value of x for which the Prob{ X < x } = q:
2nd-Vars, invNorm(q, μ, σ), yields the x.

To find the value of x for which Prob{ X > x } = q:
2nd-Vars, invNorm(1 – q, μ, σ), yields the x.