Expanded Study Guide: Mathematical Foundations for Machine Learning

Below is a **detailed breakdown** of each section and solution, with step-by-step explanations and practical examples to ensure clarity and depth.

---

1. Linear Systems and Their Solutions

Why Does a Linear System $Ax = b$ Have Only No, One, or Infinite Solutions?

Key Idea

A linear system can be represented as:

$$Ax = b,$$

where:

- A: Coefficient matrix of size $m \times n$.
- x: Unknown variable vector of size $n \times 1$.
- b: Result/output vector of size $m \times 1$.

The number of solutions depends on the relationship between:

1. The rank of the matrix A, denoted as $\text{rank}(A)$, which is the number of linearly independent rows or columns in A.
2. The rank of the augmented matrix $[A|b]$, which includes A and b as a single matrix.

Case Analysis

1. **No Solution (Inconsistent System)**:
   - If b lies outside the column space of A, there is no solution.
   - Example:

     $$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

     Substituting into:

     $$x_1 + x_2 = 3, \quad x_1 - x_2 = 2.$$

     Adding these:

     $$2x_1 = 5 \implies x_1 = 2.5.$$

     Subtracting gives:

     $$2x_2 = 1 \implies x_2 = 0.5.$$

     However, if b changes to:

     $$b = \begin{bmatrix} 3 \\ 5 \end{bmatrix},$$

     the system becomes inconsistent, as no combination of x_1, x_2 satisfies both equations.

2. **Unique Solution (Consistent and Independent)**:
   - Occurs if $\text{rank}(A) = n$, and the system is consistent.
   - Example:

     $$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

     Solve:

     $$x_1 + x_2 = 3, \quad x_1 - x_2 = 1.$$

     Add:

$$2x_1 = 4 \implies x_1 = 2, \quad x_2 = 1.$$

3. **Infinite Solutions (Dependent System)**:
   - Occurs if $\text{rank}(A) < n$, but b lies in the column space of A.
   - Example:

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

   Reduce:

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 2 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

   The system is consistent but dependent, with infinitely many solutions:

$$x_1 + x_2 = 3, \quad x_2 = t \implies x_1 = 3 - t.$$

---

2. Solving a 10 × 5 Linear System $Ax = b$ in Two Halves

Key Idea

A 10×5 system involves solving for x (a vector of size 5×1) using 10 equations. Splitting it into two systems:

1. First 5 equations: $A_1 x = b_1$, where A_1 is 5×5, and b_1 is 5×1.
2. Last 5 equations: $A_2 x = b_2$, where A_2 is 5×5, and b_2 is 5×1.

Challenges

1. **Potential Inconsistency**:
   - Splitting assumes both subproblems are independently consistent. If they are not, solving them separately will yield different or no solutions.
2. **Loss of Global Information**:
   - Solving $A_1x = b_1$ alone ignores constraints from $A_2x = b_2$.

Example

1. Full system:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \\ 26 & 27 & 28 & 29 & 30 \\ 31 & 32 & 33 & 34 & 35 \\ 36 & 37 & 38 & 39 & 40 \\ 41 & 42 & 43 & 44 & 45 \\ 46 & 47 & 48 & 49 & 50 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{bmatrix}.$$

2. Split into:

$$A_1x = b_1, \quad A_2x = b_2.$$

Solving each independently may yield different x.

---

3. Counting Operations for REF to RREF

Row Reduction

Converting A to RREF involves:

1. Creating pivot elements.
2. Eliminating all other entries in pivot columns.

Operation Complexity

1. For an $n \times n$ matrix:
   - **Gaussian Elimination**: $O(n^3)$ operations.
   - **REF to RREF**: Additional $O(n^2)$.

---

4. Consistency in Combined Systems

Given:

1. $Ax_1 = b_1, Ax_2 = b_2$.
2. Combined system: $A(x_1 + x_2) = b_1 + b_2$.

Proof:

$$A(x_1 + x_2) = Ax_1 + Ax_2 = b_1 + b_2.$$

Thus, the combined system is consistent if $b_1 + b_2$ lies in the column space of A.

---

5. Non-Zero Matrices A, B Such That $AB = 0$ and $BA \neq 0$

Key Idea

Two non-zero matrices A and B can multiply to 0 if their null spaces align. However, their product may not commute.

---

6. Proving Positive Definiteness

Definition:

A matrix $S = A^T A$ is positive definite if:

1. $x^T Sx > 0$ for all non-zero x.
2. Diagonal entries of S are non-negative.

Proof:

1. For any x:

$$x^T Sx = ||Ax||^2.$$

   Since $||Ax||^2 > 0$ for $x \neq 0$, S is positive definite.

---

7. Determinant Using Recursion

Algorithm:

1. Expand along the first row:

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{1j}),$$

where A_{1j} is the minor matrix obtained by removing the j-th column and first row.

Complexity:

- Recursive method: $O(n!)$.
- Gaussian elimination: $O(n^3)$.