A. APPLICATION OF MATRICES IN AI & DATA SCIENCE

1. Introduction

Matrices are used in AI to store, transform and process data. Images, signals, text vectors and model weights are all represented as matrices. Deep learning models like CNNs use matrix operations for feature extraction.

2. Convolution in CNNs

Convolution is an operation where a small matrix (filter) moves over an image matrix.

Each position: multiply element-wise, add the values, and store the result. This produces a feature map.

It helps detect edges, corners, textures and patterns.

3. Convolution Example (Draw by hand in notebook)

- 5×5 image matrix

- 3×3 filter

- 3×3 feature map

(Label them properly when writing.)

4. Application in AI & DS

Used in handwritten digit recognition (MNIST):

1) Input image stored as a 28×28 matrix

2) Filters extract strokes, curves, edges

3) Feature maps combine layer-by-layer

4) Final prediction 0–9

Also used in object detection, medical imaging and face recognition.

5. Advantages

• Efficient computation

• Works well on GPUs

• Extracts meaningful features

• Suitable for deep learning

6. Limitations

• High memory usage

• Expensive computation for deep networks

• Numerical instability in some operations

Mitigation: Sparse matrices, normalization, PCA/SVD.

B. APPLICATION OF PARTIAL DERIVATIVES IN AI & DATA SCIENCE

1. Role of Partial Derivatives

Machine learning models learn by minimizing a loss function. Partial derivatives indicate how loss changes with respect to each parameter. Used in gradient descent.

2. Linear Regression Cost Function

Model: $y = mx + c$

Loss function: $L = (1/2m)\ \Sigma\ (y - (mx + c))^2$

3. Partial Derivatives

$\partial L/\partial m = -(1/m)\ \Sigma\ x\ (y - y\_hat)$

$\partial L/\partial c = -(1/m)\ \Sigma\ (y - y\_hat)$

These guide how slope and intercept should adjust.

4. Gradient Update Example (Draw table manually)

Dataset: (1,2), (2,3), (3,5)

Steps:

1) Predict using initial m=0, c=0

2) Compute loss

3) Compute gradients

4) Update parameters using:

$m = m - \alpha(\partial L/\partial m)$

$c = c - \alpha(\partial L/\partial c)$

5. Applications

Used in price prediction, trend forecasting, fraud detection and recommendation systems.

6. Advantages

• Efficient optimization

• Supports large datasets

• Backbone of neural network training (backpropagation)

7. Limitations

• Can get stuck in local minima

• Slow convergence

• Sensitive to learning rate

Mitigation: Adam, RMSProp, learning rate schedules, normalization.

Conclusion

Matrices support structured data representation and feature extraction in AI.

Partial derivatives guide optimization during model training.

Together, they form the mathematical foundation of modern AI and Data Science.