About the MCC

N.B., the notation here applies only to this section.

1 MCC for the Binary Classification Extension

In multi-class problems, the binary classification extension treats each class as a binary classification problem (one-vs-rest). MCC for each class i is calculated individually, where the rest of the classes are treated as negative. The MCC for each class is given by:

$$\text{MCC}_i = \frac{\text{TP}_i \cdot \text{TN}_i - \text{FP}_i \cdot \text{FN}_i}{\sqrt{(\text{TP}_i + \text{FP}_i)(\text{TP}_i + \text{FN}_i)(\text{TN}_i + \text{FP}_i)(\text{TN}_i + \text{FN}_i)}}$$

where:

- TP_i represents the True Positives for class i.

- TN_i represents the True Negatives for class i.

- FP_i represents the False Positives for class i.

- FN_i represents the False Negatives for class i.

The overall MCC for the multi-class problem is then computed by averaging or using a weighted average of each class's MCC:

$$\text{MCC}_{\text{overall}} = \frac{1}{N} \sum_{i=1}^{N} \text{MCC}_i$$

where N is the total number of classes.

The MATLAB code for Binary Classification Extension is as follow:

```matlab
function mcc = mcc_bin_ext(confMat)
    numClasses = size(confMat, 1);
    TP = zeros(numClasses, 1);
    TN = zeros(numClasses, 1);
    FP = zeros(numClasses, 1);
    FN = zeros(numClasses, 1);
    for i = 1:numClasses
        TP(i) = confMat(i, i);
        FP(i) = sum(confMat(:, i)) - TP(i);
        FN(i) = sum(confMat(i, :)) - TP(i);
        TN(i) = sum(confMat(:)) - (TP(i) + FP(i) + FN(i));
    end
    numerator = sum(TP .* TN) - sum(FP .* FN);
    denominator = sqrt(sum((TP + FP) .* (TP + FN)) * sum((TN + FP) .*
        (TN + FN)));
```

```
    if denominator == 0
        mcc = 0;
    else
        mcc = numerator / denominator;
    end
end
```

2 MCC for the Direct Multi-Class Classification

The direct method computes the MCC for the entire confusion matrix without decomposing it into binary classification problems. The formula is expressed as:

$$\text{MCC} = \frac{c \cdot s_1 - s_2}{\sqrt{(c^2 - s_3) \cdot (c^2 - s_4)}}$$

where:

- $c = \sum_{i,j} \text{ConfMat}_{ij}$ is the total sum of the confusion matrix (all elements combined).

- $s_1 = \sum_k \text{ConfMat}_{kk}$ represents the sum of the diagonal elements (correctly classified instances).

- $s_2 = \sum_k \text{row\_sum}_k \cdot \text{col\_sum}_k$, where row_sum_k is the sum of elements in row k, and col_sum_k is the sum of elements in column k.

- $s_3 = \sum_k \text{row\_sum}_k^2$, representing the sum of squared row sums.

- $s_4 = \sum_k \text{col\_sum}_k^2$, representing the sum of squared column sums.

The MATLAB code for Direct Multi-Class MCC is as follow:
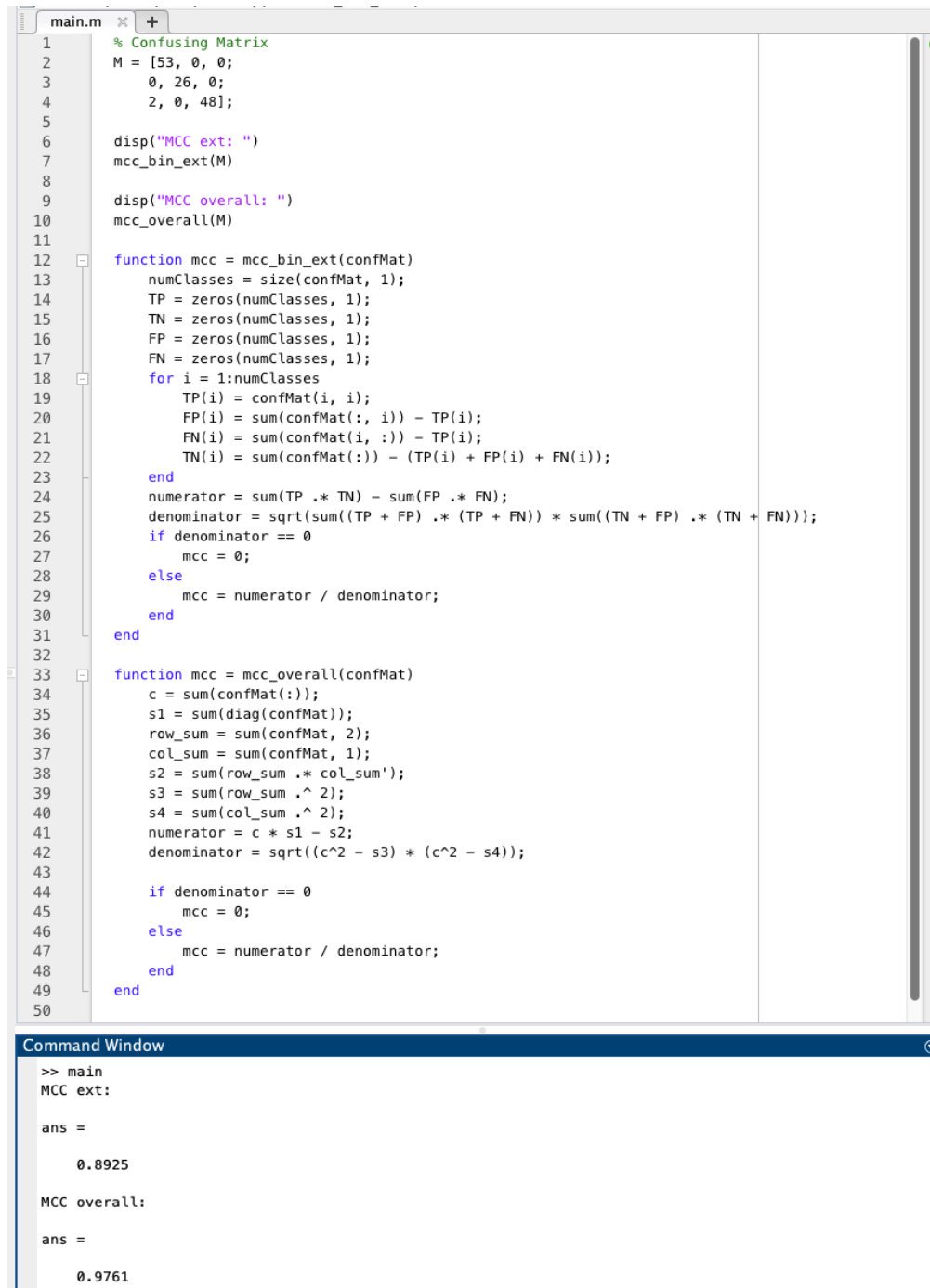
```
function mcc = mcc_overall(confMat)
    c = sum(confMat(:));
    s1 = sum(diag(confMat));
    row_sum = sum(confMat, 2);
    col_sum = sum(confMat, 1);
    s2 = sum(row_sum .* col_sum');
    s3 = sum(row_sum .^ 2);
    s4 = sum(col_sum .^ 2);
    numerator = c * s1 - s2;
    denominator = sqrt((c^2 - s3) * (c^2 - s4));

    if denominator == 0
        mcc = 0;
    else
        mcc = numerator / denominator;
    end
end
```

3 Code

The "main.m" contains the functions and instance.

```matlab
% Confusing Matrix
M = [53, 0, 0;
      0, 26, 0;
      2, 0, 48];

disp("MCC ext: ")
mcc_bin_ext(M)

disp("MCC overall: ")
mcc_overall(M)

function mcc = mcc_bin_ext(confMat)
    numClasses = size(confMat, 1);
    TP = zeros(numClasses, 1);
    TN = zeros(numClasses, 1);
    FP = zeros(numClasses, 1);
    FN = zeros(numClasses, 1);
    for i = 1:numClasses
        TP(i) = confMat(i, i);
        FP(i) = sum(confMat(:, i)) - TP(i);
        FN(i) = sum(confMat(i, :)) - TP(i);
        TN(i) = sum(confMat(:)) - (TP(i) + FP(i) + FN(i));
    end
    numerator = sum(TP .* TN) - sum(FP .* FN);
    denominator = sqrt(sum((TP + FP) .* (TP + FN)) * sum((TN + FP) .* (TN + FN)));
    if denominator == 0
        mcc = 0;
    else
        mcc = numerator / denominator;
    end
end

function mcc = mcc_overall(confMat)
    c = sum(confMat(:));
    s1 = sum(diag(confMat));
    row_sum = sum(confMat, 2);
    col_sum = sum(confMat, 1);
    s2 = sum(row_sum .* col_sum');
    s3 = sum(row_sum .^ 2);
    s4 = sum(col_sum .^ 2);
    numerator = c * s1 - s2;
    denominator = sqrt((c^2 - s3) * (c^2 - s4));

    if denominator == 0
        mcc = 0;
    else
        mcc = numerator / denominator;
    end
end
```

```
Command Window
>> main
MCC ext:

ans =

    0.8925

MCC overall:

ans =

    0.9761
```

3