OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• **Laplacian eigenfunctions and neural networks:...** The detailed numerical example presented on this page describes the computational steps involved to compute the value of the performance metric from the previous page. First, we use the ground truth classification matrix G – with only correct classification outcomes – to define the "inverse occurrence weights" $1/s_{cl}$, where a class cl arises s_{cl} times. The diagonal scale matrix S with diagonal entries is used to multiply the result matrix R obtained by the actual classification system, generating the $S \bullet R$ matrix values.

$$\begin{array}{cc|ccc} & & & \xrightarrow{cl2} & \\ & & 0 & 1 & 2 \\ \hline cl1 \downarrow & 0 & 20 & & \\ & 1 & & 4 & \\ & 2 & & & 8 \end{array} \qquad G$$

$s_0 = 20,\ 1/s_0 = 1/20$
$s_1 = 4,\ 1/s_1 = 1/4$
$s_2 = 8,\ 1/s_2 = 1/8$

$$\underbrace{\begin{bmatrix} 1/20 & & \\ & 1/4 & \\ & & 1/8 \end{bmatrix}}_{S} \bullet \underbrace{\begin{bmatrix} 15 & 3 & 2 \\ 1 & 3 & 0 \\ 4 & 1 & 3 \end{bmatrix}}_{R} = \underbrace{\begin{bmatrix} 3/4 & 3/20 & 1/10 \\ 1/4 & 3/4 & 0 \\ 1/2 & 1/8 & 3/8 \end{bmatrix}}_{S \bullet R}$$

$$\underbrace{\begin{bmatrix} 1 & -1 & -2 \\ -2 & 1 & -1 \\ -4 & -2 & 1 \end{bmatrix}}_{W \text{ (weights)}} \circ \underbrace{\begin{bmatrix} 3/4 & 3/20 & 1/10 \\ 1/4 & 3/4 & 0 \\ 1/2 & 1/8 & 3/8 \end{bmatrix}}_{S \bullet R} = \underbrace{\begin{bmatrix} 3/4 & -3/20 & -1/5 \\ -1/2 & 3/4 & 0 \\ -2 & -1/4 & 3/8 \end{bmatrix}}_{W \circ (S \bullet R) = M_P = \langle m_{cl1,cl2} \rangle}$$

$$\Rightarrow \underline{\underline{P}} = \sum_{cl1=0}^{2} \sum_{cl2=0}^{2} m_{cl1,cl2}$$

$$= 3/4 - 3/20 - 1/5 - 1/2 + 3/4 - 0 - 2 - 1/4 + 3/8$$

$$\underline{\underline{= -49/40}}$$

•••

OBJECT AND MATERIALS EIGENFUNCTIONS - Cont'd.

- **Laplacian eigenfunctions and neural networks:...** The weight matrix W contains the values to be used to multiply the $S \cdot R$ values; the weight in position $(cl1, cl2)$ is the specific factor to be multiplied with the $S \cdot R$ value in the same position. This operation is denoted by '$\circ$' and produces the final performance matrix $M_P = W \circ (S \cdot R)$. The value of the overall performance metric P is the sum of the values of the individual M_P matrix values, i.e., $P = \sum_{M_P} m_{cl1,cl2}$, where $M_P = (m_{cl1,cl2})$.

- **Note.** Of special interest is the best (worst) matrix M_P:

$$M_P^{best} = \begin{bmatrix} 1 & -1 & -2 \\ -2 & 1 & -1 \\ -4 & -2 & 1 \end{bmatrix} \circ \left(\begin{bmatrix} 1/20 & & \\ & 1/4 & \\ & & 1/8 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 20 & & \\ & 4 & \\ & & 8 \end{bmatrix}}_{R^{best}} \right) = \underbrace{\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}}_{P=3}.$$

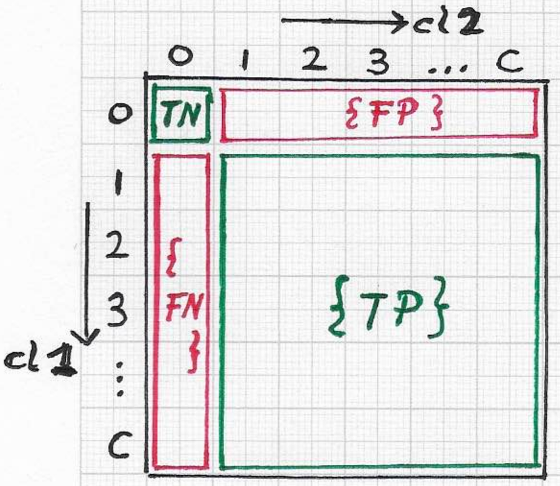
$$M_P^{worst} = \begin{bmatrix} 1 & -1 & -2 \\ -2 & 1 & -1 \\ -4 & -2 & 1 \end{bmatrix} \circ \left(\begin{bmatrix} 1/20 & & \\ & 1/4 & \\ & & 1/8 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} & & 20 \\ 4 & & \\ 8 & & \end{bmatrix}}_{R^{worst}} \right) = \underbrace{\begin{bmatrix} & & -2 \\ -2 & & \\ -4 & & \end{bmatrix}}_{P=-8}.$$

The system performs optimally when $R = G$ (top). Here, $P = 3$ is equal to $(C+1) = 3$. The worst performance results when class 0 is always classified as class C $(=2)$ and all threat classes $(1 \ldots C=2)$ are always classified as class 0 (bottom).

...

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...



	0	1
0	TN	FP
1	FN	TP

Correspondence between multi-class result outcome groups (top) and simple two-class/binary result outcome types (bottom).

One can adopt the viewpoint that mis-classifications of threat materials to incorrect threat materials are still to be viewed as TP outcomes – since a threat was detected. Thus, one can also understand the top table as a binary classification table as shown in the bottom table.

The complex, intricate methodology for classification performance characterization described applies to a multi-class classification setting. The multi-class setting is sketched in the left figure (top). A given unclassified material of (real) class cl1 is classified by the system as a class-cl2 material. On a high level, even for this (C+1)-by-(C+1) table we can essentially consider four outcome types: TN, TP, FN and FP. This viewpoint can be justified: Classifying any (real) threat material (1,...C) as a threat material (even the wrong class) is still a threat detection, and thus a TP outcome. Similarly, all classifications of a threat material as a class-0 material is an FN outcome; and classifying a class-0 material of any threat class is an FP outcome. !

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

- Laplacian eigenfunctions and neural networks:... Many classification performance metrics exist for two-class/binary classification. Some of these metrics are reviewed here. Most of the commonly used traditional, standard performance metrics capture a specific, important aspect of a classification system or for a certain performance focus. Unfortunately, many of these established **metrics** cannot be viewed independently, since their "high-level statistical nature" hides various relevant aspects of overall performance. We provide a partial list of such metrics next.

- **PRECISION** is defined as

$$prec = TP/(TP+FP) .$$

- **RECALL** is defined as

$$rec = TP/(TP+FN) .$$

Both metrics attempt to describe the number of TP classification outcomes relative to the sum of two other classification outcome numbers. Precision is also called POSITIVE PREDICTIVE VALUE (PPV), and recall is also known as SENSITIVITY.

- The F-MEASURE combines precision and recall:

$$F = 2 \cdot prec \cdot rec / (prec + rec) .$$

...

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks: ...

We consider an example. The two tables (left) show the detailed numbers of all possible 9 outcome results of a 3-class classification problem. Clearly, there is a loss of information when "condensing" the 9-outcome to the derived 4-outcome binary outcome table.

cl1 ↓ \ cl2 →	0	1	2
0	15	3	2
1	1	3	0
2	4	1	3

Full 3-by-3 outcome table

TN 15	FP 5
FN 5	TP 7

Derived 2-by-2 binary table

For this example, one obtains:

$$\text{prec} = TP/(TP+FP) = 7/12 \, .$$

$$\text{rec} = TP/(TP+FN) = 7/12 \, .$$

$$F = 2 \cdot \text{prec} \cdot \text{rec} / (\text{prec} + \text{rec}) = 7/12 \, .$$

It is evident that these three performance values "hide" details about classification behavior – as all values are 7/12. (**Note.** The optimal binary table for this example has the values $TN = 20$, $TP = 12$, $FN = 0$ and $FP = 0$. Thus, the value of the three binary metrics is $\text{prec} = \text{rec} = F = 1$.)

Additional metrics often used include the following:

- **ACCURACY** is defined as

$$\text{acc} = (TP+TN)/(P+N) \, ,$$

where P (N) is the number of actual positive (negative) data. SENSITIVITY $= TP/P$ and SPECIFICITY $= TN/N$ define the normalized metric

- **BALANCED ACCURACY**, defined as $\frac{1}{2}(TP/P + TN/N)$.

...