■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks: ... The table on p. 4 (11/17/2022) captures detailed statistical information concerning classification performance. It contains the information that summarizes how many times a material of (ground truth) class cl, $cl \in \{0, 1, \ldots, C\}$, is classified by the system as class $\overline{cl}$, $\overline{cl} \in \{0, 1, \ldots, C\}$. The general form of such a table is shown here (left). This table effectively is a square $(C+1)$-by-$(C+1)$ "result matrix" R with values $r_{cl,\overline{cl}}$. The R value $r_{cl,\overline{cl}}$ defines how often materials belonging to class cl are classified as class-$\overline{cl}$ materials.

cl \ $\overline{cl}$	0	1	...	C
0	$r_{0,0}$	$r_{0,1}$	...	$r_{0,C}$
1	$r_{1,0}$	$r_{1,1}$	...	$r_{1,C}$
⋮			⋮	
C	$r_{C,0}$	$r_{C,1}$	...	$r_{C,C}$

In the context of processing and classifying streamed data, this R matrix is continually changing. The R matrix representing our results from classifying 32 materials (p. 3, 11/17/2022) is included here (left). The "optimal" result matrix is the diagonal matrix D (shown below), where only the diagonal matrix values $r_{cl,cl}$ are different from 0. A sequence of optimal D matrices – e.g., $D_1, \ldots, D_{32}$ – produces ratios $r_{0,0} : r_{1,1} : \ldots : r_{C,C}$ that capture and converge to the occurrence ratio of class-0, ..., class-C materials.

$$R = \begin{bmatrix} 15 & 3 & 2 \\ 1 & 3 & 0 \\ 4 & 1 & 3 \end{bmatrix}$$

$$D = \begin{bmatrix} 20 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

...

Stratovan

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... Considering the specific exemplary matrices **R** and **D** from the previous page, one can ask the following questions:

1) Which "transformation matrix" **T** must be applied to **D** to obtain the result matrix **R**?

2) Which "transformation matrix" $\tilde{T}$ must be applied to the result matrix **R** to obtain **D**?

The answer of the first question is the solution of $TD=R$, and the answer of the second question is the solution of $\tilde{T}R=D$. Thus, we solve these equations for T and $\tilde{T}$, respectively: $T=RD^{-1}$ and $\tilde{T}=DR^{-1}$.

$$D=\begin{bmatrix}20&0&0\\0&4&0\\0&0&8\end{bmatrix} \Rightarrow D^{-1}=\begin{bmatrix}1/20&0&0\\0&1/4&0\\0&0&1/8\end{bmatrix} \Rightarrow T=\begin{bmatrix}15&3&2\\1&3&0\\4&1&3\end{bmatrix}D^{-1}=\begin{bmatrix}3/4&3/4&1/4\\1/20&3/4&0\\1/5&1/4&3/8\end{bmatrix}$$

$$R=\begin{bmatrix}15&3&2\\1&3&0\\4&1&3\end{bmatrix} \Rightarrow R^{-1}=1/104\begin{bmatrix}45&-35&-30\\-3&37&2\\-22&-6&84\end{bmatrix} \Rightarrow \tilde{T}=\begin{bmatrix}20&0&0\\0&4&0\\0&0&8\end{bmatrix}R^{-1}=1/26\begin{bmatrix}45&-35&-30\\-3&37&2\\-22&-6&-84\end{bmatrix}$$

Why are these matrices of interest? What is the meaning of these matrices? How are these matrices related to our classification goals, and how could one potentially use the "information content" of these matrices to further improve classification results?

• ONE CAN INTERPRET $TD=R$ AS FOLLOWS: "The observed result **R** is the consequence of a transformation **T** acting on the optimal diagonal matrix **D**, representing the best-possible classification result."

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... SIMILARLY, ONE CAN INTERPRET $\tilde{T}R = D$ AS FOLLOWS: "The optimal diagonal matrix D can be obtained by having a transformation matrix $\tilde{T}$ act on the observed result R."

⇒ The matrices T and $\tilde{T}$ are of interest as they characterize the classification system's deviation from the best-possible performance. The matrices T and $\tilde{T}$ define, in a numerically exact manner, how one has to manipulate the observed classification results captured statistically in the result matrix R to obtain the best-possible, perfect diagonal matrix D. In summary: THE TRAINED, TESTED AND VALIDATED CLASSIFICATION SYSTEM HAS A STATISTICAL PERFORMANCE CHARACTERISTIC CAPTURED IN THE MATRICES R, D, T AND $\tilde{T}$ – WHEN CONSIDERING A FINITE NUMBER OF PROCESSED STREAM DATA. CONCERNING INCOMING, NOT-YET-CLASSIFIED STREAM DATA, ONE SHOULD USE THESE MATRICES TO INFLUENCE THE SYSTEM'S CLASSIFCATION DECISIONS SUCH THAT CLASSIFICATION RESULTS ARE OPTIMIZED.

...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

- Laplacian eigenfunctions and neural networks:... Since it not intuitive or straightforward how to utilize the R, D, T and $\tilde{T}$ matrices to improve classification performance, we consider additional simple examples. First, we consider a binary classification scenario, where only classes 0 and 1 are possible:

TRUE stream: 00001111
Classification as: 11110000

$$\Rightarrow R = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}.$$

The classification system's behavior can be simply described: "Class 0 is always mis-classified as class 1, and class 1 is always mis-classified as class 0." Thus, the (initial) classification system's output 0 (1) must always be changed to the (final) output 1 (0). The perfectly corrected (final) classification output stream is: 00001111. One can view an initial classification as a "prediction" and a final classification as a "correction." The T and $\tilde{T}$ matrices for this scenario are the matrices

$$T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

For example, our goal could be to "apply $\tilde{T}$ to a prediction to obtain the correction." We must devise a way that allows us to apply the 2-by-2 matrix $\tilde{T}$ to an initially predicted class, i.e., a single integer. FOR THIS PURPOSE, WE USE THE CONCEPT OF CLASS-FRACTION DATA.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... WE DEFINE CLASS-FRACTION DATA FOR A SET OF BASIC/ELEMENTARY CLASSES AS FOLLOWS: CLASS-FRACTION ($\mathcal{F}$) DATA ARE "CONVEX COMBINATIONS" OF BASIC CLASSES $0, 1, \ldots, C$ ($cl_0, cl_1, \ldots, cl_C$), I.E., $\mathcal{F} = w_0\, cl_0 + w_1\, cl_1 + \ldots + w_C\, cl_C$, WHERE $w_i \geq 0$ AND $\sum_i w_i = 1$. AN $\mathcal{F}$ DATUM IS IDENTIFIED WITH ITS WEIGHT TUPLE $(w_0, \ldots, w_C)$. This definition makes it possible to view class 0 as tuple $(1,0)$ and class 1 as $(0,1)$ – in our binary classification context. We can now apply the 2-by-2 matrix $\tilde{T}$ to this 2D representation. We apply $\tilde{T}$ to class 0 and class 1:

$$\tilde{T} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \tilde{T} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Thus, the application of $\tilde{T}$ to the initial predictions $(1,0)$ and $(0,1)$ produces the corrections $(0,1)$ and $(1,0)$, respectively. Thus, the final classifications generated by the matrix calculations are corrections that are perfect, optimal. We now consider a slightly more complicated example:

TRUE stream: 00001111
Classification as: 00111111

$$\Rightarrow R = \begin{bmatrix} 2 & 2 \\ 0 & 4 \end{bmatrix}, \quad R^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ 0 & \frac{1}{4} \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix},$$

$$D^{-1} = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix}.$$

...