# 22) Style Transfer

## Gram Matrix

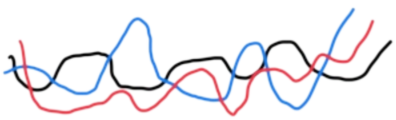
- Gram Matrix is ② without $n-1$ data-scaling factor & mean centering

① Correlation : contains correlation coefficients b/t all pairs of the dataset features

- Formula: $r_{x,y} = \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \,/\, \sqrt{\sum_{i=1}^{n}(x_i-x)^2(y_i-y)^2}$

② Covariance : diagonal = var(channel), non-diagonal = cov(channel A, channel B)

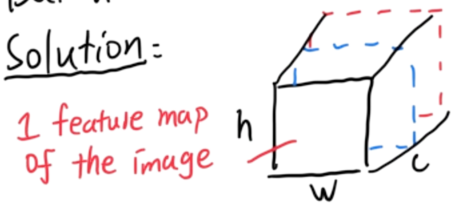
- Formula: $C_{x,y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$
- Interpretation: represents all linear interactions across feature cols (for PCA)
  captures abstract patterns ('texture' / 'style') of the image



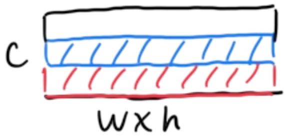
$$\begin{array}{c|c} & 1 \quad\; 2 \quad\; 3 \\ \begin{matrix}1\\2\\3\end{matrix} & \begin{bmatrix} 3.16 & 2.79 & 4.13 \\ 2.79 & 11.79 & 4.32 \\ 4.13 & 4.32 & 9.31 \end{bmatrix} \end{array} \xrightarrow[\text{to } [-1,1]]{\text{Normalize}} \begin{array}{c|c} & 1 \quad\; 2 \quad\; 3 \\ \begin{matrix}1\\2\\3\end{matrix} & \begin{bmatrix} 1 & 0.51 & 0.67 \\ 0.51 & 1 & 0.41 \\ 0.67 & 0.41 & 1 \end{bmatrix} \end{array}$$

① ②

- Gram Matrix: $G = \bar{X}^T\bar{X}$, where $\bar{X}$ is pixels of 1 feature map
- But how to compute $G$ for a convolution layer (>1 feature maps)?
- Solution:




Flatten $\Longrightarrow$




$= \bar{X}$ : $G$.shape = (c, c)

## Style Transfer Algorithm

- Overview: Content Image + Style Image = Target Image


I = Image
G = Gram matrix

- Steps: ① Initialize random target I ② Pick pretrain CNN network
  ③ Match content in early layers ($L_c$ = MSE(target I, content I))
  ④ Match style in convolution layers ($L_s$ = MSE($\alpha_i$(target G, style G))
  ⑤ Backprop target I on Loss = $L_c + \beta L_s$

style weight for conv layer $i$ decreasing as $i$ gets higher

style scaling = 1e6