

1 Effective Projective Field Size (One dimensional scenario)

What follows formalizes the effective projective field growth for the one dimensional scenario with n convolution layers (assuming certain conditions).

The input is $u(t)$ where

$$u(t) = \begin{cases} 1, & t = 0 \\ 0, & t \neq 0 \end{cases} \quad (1)$$

and $t = 0, 1, -1, 2, -2, \dots$ indexes the input pixels.

Each layer has the **same kernel** $v(t)$ of size k . The kernel signal can be formally defined as

$$v(t) = \sum_{m=0}^{k-1} w(m)\delta(t-m) \quad (2)$$

where $w(m)$ is the weight for the m th pixel in the kernel. Without losing generality, we can assume the weights are normalized, i.e. $\sum_m w(m) = 1$. The output signal of the n th layer $o(t)$ is simply $o = u * v * \dots * v$, convolving u with n such v 's. To compute the convolution, we can use the Discrete Time Fourier Transform to convert the signals into the Fourier domain, and obtain

$$U(\omega) = \sum_{t=-\infty}^{\infty} u(t)e^{-j\omega t} = 1, \quad V(\omega) = \sum_{t=-\infty}^{\infty} v(t)e^{-j\omega t} = \sum_{m=0}^{k-1} w(m)e^{-j\omega t} \quad (3)$$

Applying the convolution theorem, we have the Fourier transform of o is

$$\mathcal{F}(o) = \mathcal{F}(u * v * \dots * v)(\omega) = U(\omega) \cdot V(\omega)^n = \left(\sum_{m=0}^{k-1} w(m)e^{-j\omega t} \right)^n \quad (4)$$

With inverse Fourier transform

$$o(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{m=0}^{k-1} w(m)e^{-j\omega t} \right)^n e^{j\omega t} d\omega \quad (5)$$

The space domain signal $o(t)$ is given by the coefficients of $e^{-j\omega t}$. These coefficients turn out to be well studied in the combinatorics literature [1]. It can be shown that if $\sum_{\mathbf{m}} \mathbf{w}(\mathbf{m}) = \mathbf{1}$ and $\mathbf{w}(\mathbf{m}) \geq \mathbf{0} \forall \mathbf{m}$, then

$$o(t) = p(S_n = t), \text{ where } S_n = \sum_{i=1}^n X_i \text{ and } p(X_i = m) = w(m) \quad (6)$$

From the central limit theorem, as $n \rightarrow \infty$, $\sqrt{n}(\frac{1}{n}S_n - \mathbb{E}[X]) \sim \mathcal{N}(0, \text{Var}[X])$ and $S_n \sim \mathcal{N}(n\mathbb{E}[X], n\text{Var}[X])$. As $o(t) = p(S_n = t)$, $o(t)$ also has a Gaussian

shape with

$$\mathbb{E}[S_n] = n \sum_{m=0}^{k-1} mw(m), \text{Var}[S_n] = n \left(\sum_{m=0}^{k-1} m^2 w(m) - \left(\sum_{m=0}^{k-1} mw(m) \right)^2 \right) \quad (7)$$

This indicates that $o(t)$ decays from the center of the projective field squared exponentially according to the Gaussian distribution. As the rate of decay is related to the variance of the Gaussian and assuming the size of the effective projective field is one standard deviation, the size can be expressed as

$$\sqrt{\text{Var}[S_n]} = \sqrt{n \text{Var}[X_i]} = O(\sqrt{n}) \quad (8)$$

On the other hand stacking more convolution layers would grow the theoretical projective field linearly. But the effective projective field size is shrinking at a rate of $O(1/\sqrt{n})$.

References

- [1] Steffen Eger. Restricted weighted integer compositions and extended binomial coefficients. *J. Integer Seq*, 16(13.1):3, 2013.