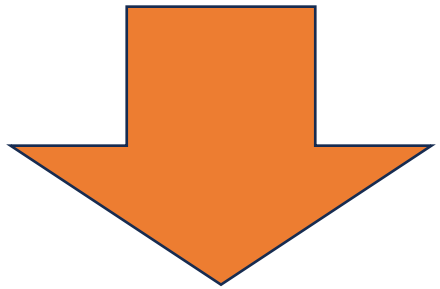
Neural Collapse in Quantized Neural Networks

Gábor Balázs Régely, ABX5LX

Introduction

Black box behavior

Heavy compute requirements



Interplay between neural collapse and quantization

Neural Collapse

- Simple **geometric form** of the features and of the classifier
- During the **terminal phase of training,** when **zero error** is achieved
- Four interconnected phenomena
- Introduced by Papyan et al. in 2020, investigated by Kothapalli

NC1 - Within-class variability collapse

- Within-class variation of activations becomes negligible as they collapse to their class-means

- The within-class covariance matrix approaches zero:

$$\sum S_w \rightarrow 0$$

NC2 - Class mean convergence to simplex ETF

- Class-means converge to an equiangular tight frame (after centering)
- Maximizing pairwise angles and distances

$$\mu_c = \frac{< \mu_c^t, \mu_{c'}^t >}{|||\mu_c^t|| \cdot ||\mu_{c'}^t||} \rightarrow \begin{cases} 1, & \text{if } c = c' \\ -\frac{1}{C-1}, & \text{if } c \neq c' \end{cases}$$

NC3 - Self-dual alignment

- Columns of the last layer linear classifier also form and converge to the simplex ETF (up to rescaling) of the penultimate layer features

$$\left\| \frac{\mu_c^t}{\|\mu_c^t\|} - \frac{w_c^t}{\|w_c^t\|} \right\| \to 0$$

NC4 - Nearest class center classification

- Last-layer classifier acts with the nearest class mean decision rule on the penultimate layer features

$$\hat{c}^t = \arg\min_{c'} ||h(x) - \mu^t_{c'}||$$

Quantization

- Comprehensive survey by Gholami et al., detailing the main quantization approaches

- EfQAT by Ashkboos et al. is a framework for QAT that reduced the computational overhead while maintaining accuracy

Computer Vision Experiments

- Metrics of Papyan et al.

- Convolutional neural networks
  - Custom CNNs
  - ResNet-18
  - MobileNetV3
  - Base ConvNeXt variants and their customized versions

- MNIST, CIFAR-10, CIFAR-100

ResNet-18, MNIST Example

≈ 4 hours

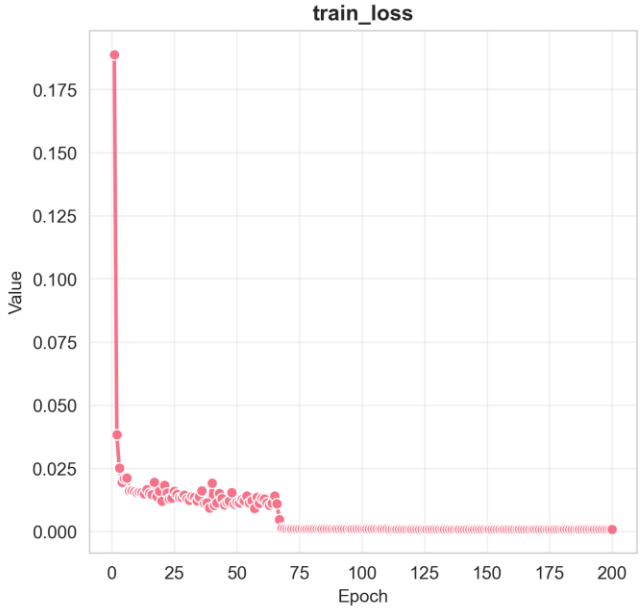
SGD with LR scheduling

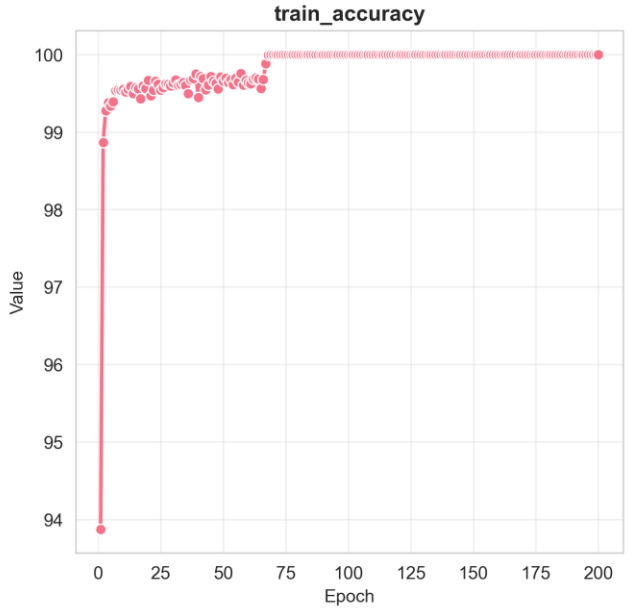
0.9 momentum

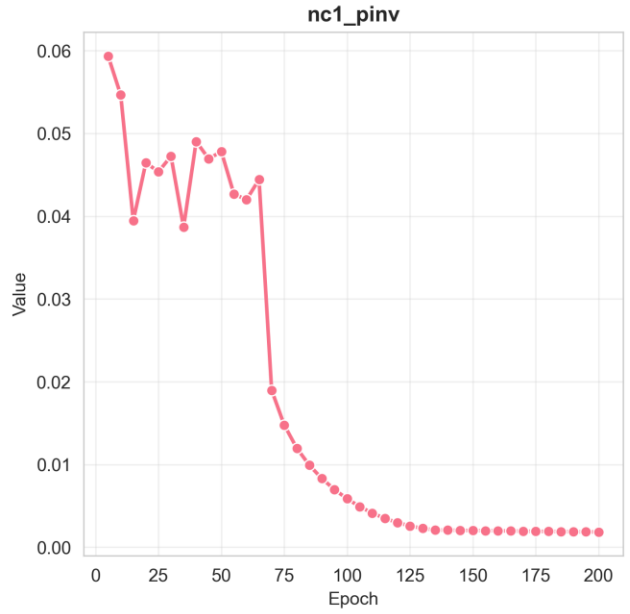
5e-4 weight decay

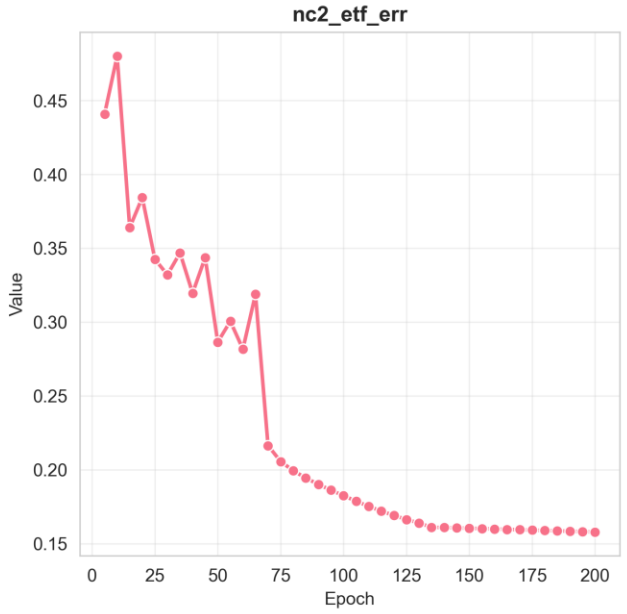
128 batch size

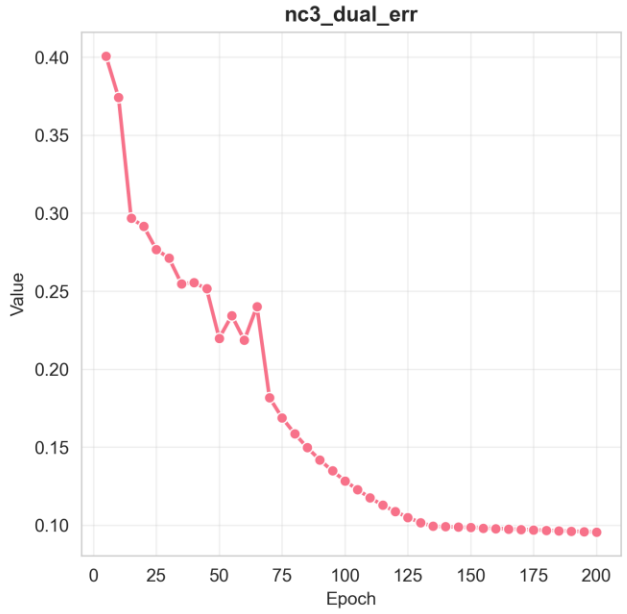
float32 precision

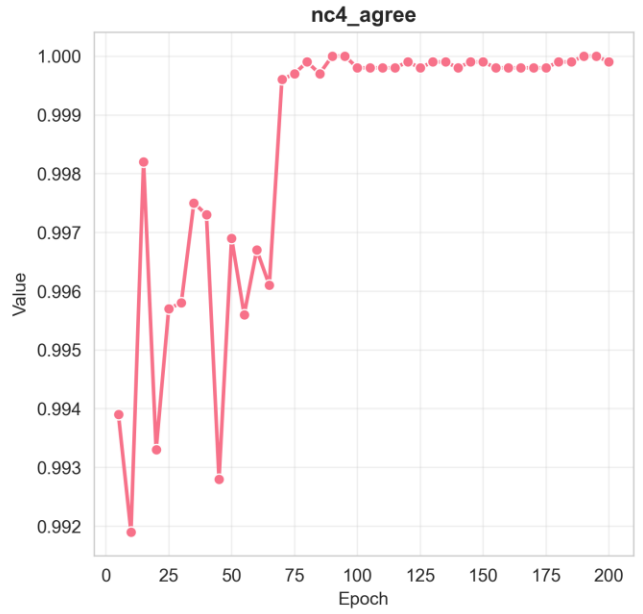












Linguistic Collapse

Conditions causing NC in NLP are different then in CV

- LMs are typically undertrained
- Classes are imbalanced
- Class numbers exceed the embedding dimension
- Contexts can be ambiguous to the next token prediction

Lut et al. showed that minimizing NC3 improves fairness scores

Language Modeling

- Empirical metrics of Wu and Papyan (2024)
- Character modeling on Shakespeare with nanoGPT
- Subsampled, class-balanced TinyStories dataset
- A simple, custom GPT model

Hyperparameters for NLP

Parameter count	1 – 10M
Loss function	Cross-entropy
Activation function	GELU
Minibatch size	32, 64, 128
Weight initialization	Normal or Kaiming
Optimizer	AdamW
Learning rate	1e-6 – 1e-3
Momentum	0.9
Betas	(0.9, 0.95)
Weight decay	0.01, 5e-4
Learning rate scheduling	Cosine, exponential, multistep
Quantization	32-bit, 16-bit floating point
Dropout; l_1, l_2 weighting	0
Layer normalization	True (norm first)

NC considerations

- **Weight initialization**
  - D'Angelo et al. studied the loss stabilizer effect of weight decay
  - Shown that it does not prevent NN from fully memorizing training data, as I also confirmed experimentally
- **Weight tying**
  - Press et al. recommended sharing the input token embeddings directly with the output classifier layer weights to reduce parameters
  - For NC this acts as a severe structural constraint

Future Work

- Improving theoretical and practical foundation

- Trials with quantization configurations
- Different Language Models

- Interaction of NC with post-training techniques
- Benchmarking downstream performance on different LLM metrics

Thank you for your attention!

AI Usage

- During the project, I trained CV and NLP AI models on publicly available datasets.
- Gemini 3.0 Pro, 3.1 Pro: Correcting LaTeX page structure and citation formatting.
- Claude Sonnet 4.5, 4.6: Generating Python functions for visualization, performance improvements on various functions.