

# 学習の安定化 – 未学習・過学習の回避

# アジェンダ

- 過学習を抑制する方法
- 過学習が起きる理由
- 正則化とは?
- (パラメータ)正則化
  - L1正則化
  - L2正則化
  - Elastic Net
- 正則化レイヤー
  - Dropout
- 正規化レイヤー
  - Batch正規化
  - Layer正規化
  - Instance正規化
- ハンズオン



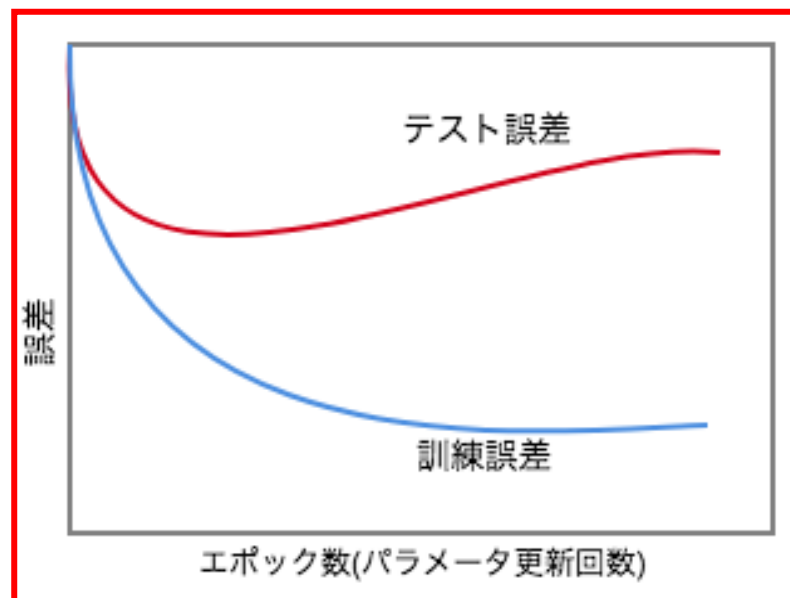
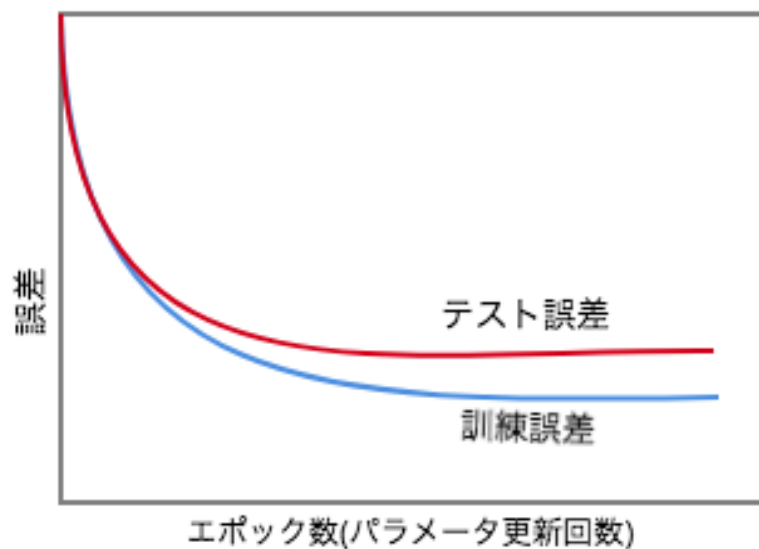
# 過学習を抑制する方法

ニューラルネットワークを前提とした機械学習において、過学習を抑制するための手法をご説明します。

(パラメータ)正則化	正則化レイヤー	正規化レイヤー
L1正則化	Dropout	Batch正規化
L2正則化		Layer正規化
Elastic Net		Instance正規化

# 過学習が起きる理由

訓練サンプルに**だけ適合**した学習をした結果。



本来の目的である汎化性能が得られない。



# 過学習が起きる理由

過学習が起きる原因の例。

- パラメータの数が多すぎる。
- パラメータの値が偏っている。
- ノードが多すぎる。
- 学習データが不足している。
- 学習データが偏っている。
- etc...

← モデルが複雑  
過じる

# 正則化とは?

ネットワークの自由度(層数、ノード数、パラメータの値、etc...)を制約すること。 Regularization。



モデルの複雑さを抑制。



訓練サンプルに対する過剰適合を抑制。



# 正則化 – L1、L2正則化

誤差に対して正則化項を加算することで、重み(パラメータ)を抑制する。

$$E_n(\mathbf{w}) + \frac{1}{p} \lambda \|x\|_p$$

誤差関数にpノルムを加える

$$\|x\|_p = \left( |x_1|^p + \dots + |x_n|^p \right)^{\frac{1}{p}}$$

pノルムの計算

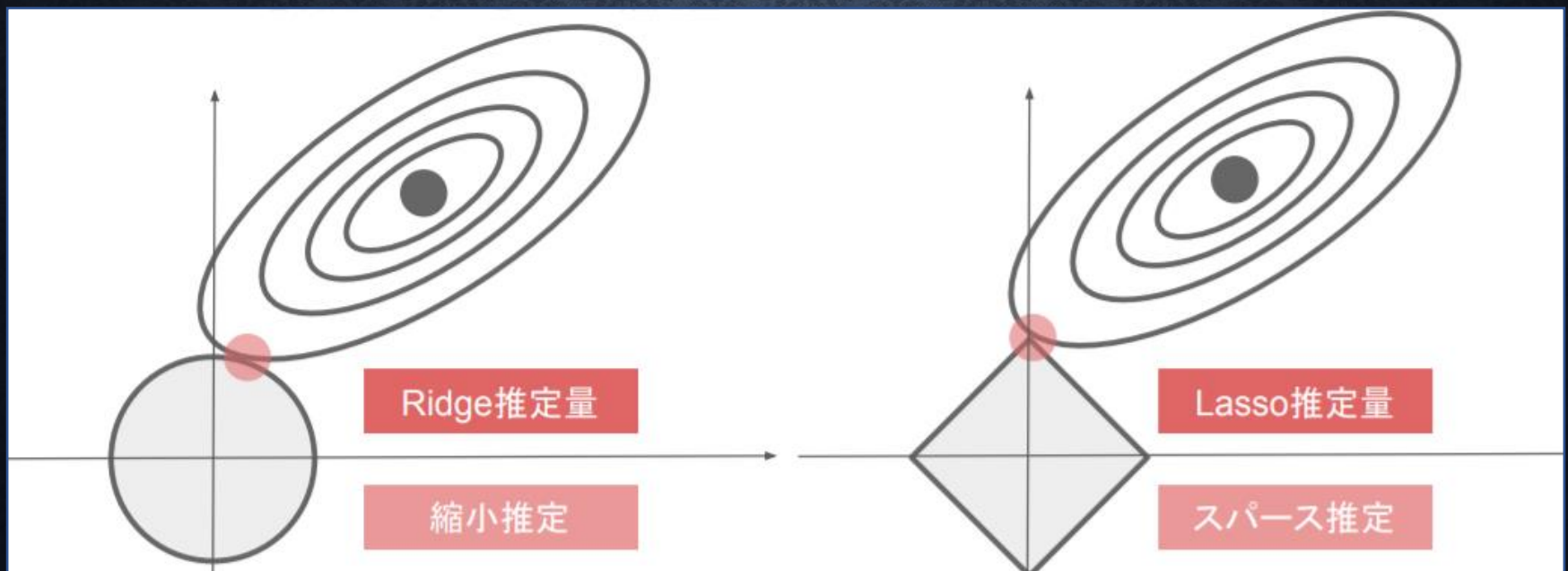
p = 1の場合、 L1正則化と呼ぶ。 Lasso回帰。  
p = 2の場合、 L2正則化と呼ぶ。 Ridge回帰。

# 正則化 – L1、L2正則化

L2正則化はパラメータが発散することを抑制、  
L1正則化はパラメータ自体を削減。

L2正則化

L1正則化





# 正則化 – Elastic Net

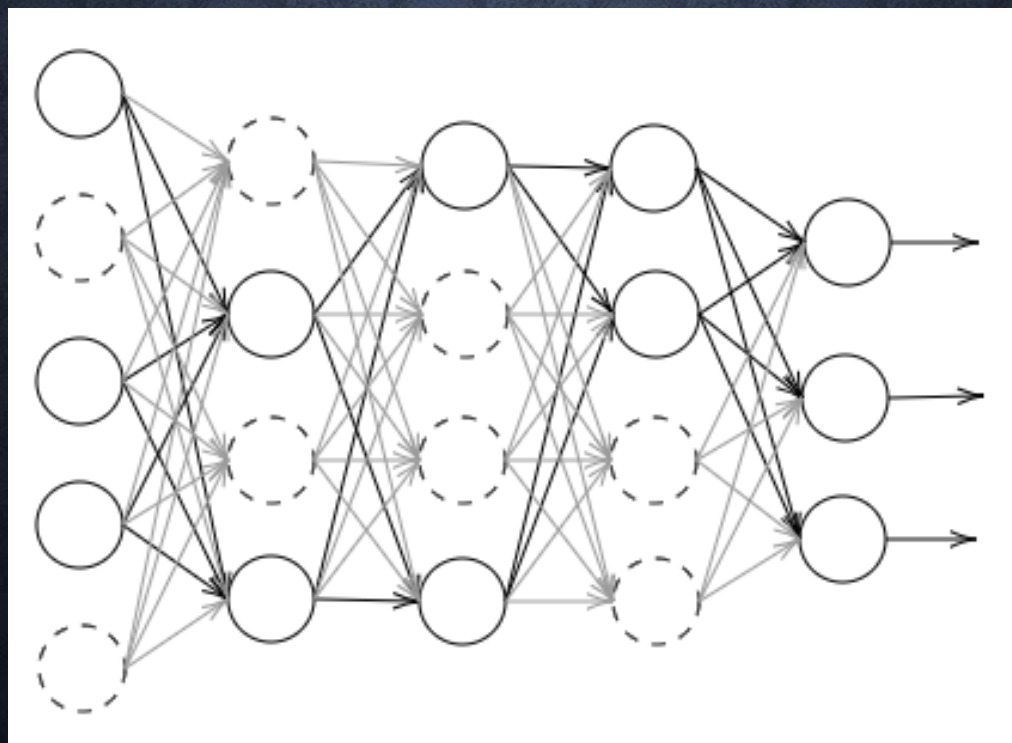
L1、L2正則化を組み合わせたもの。

$$\text{Elastic Net} = \sum_{i=1}^n \left( y^{(i)} - \hat{y}^{(i)} \right)^2 + \underbrace{\lambda_1 \sum_{j=1}^m w_j^2}_{\text{L2正則化項}} - \underbrace{\lambda_2 \sum_{j=2}^m |w_j|}_{\text{L1正則化項}}$$

$\lambda_1, \lambda_2$ はハイパーパラメータ。各正則化項の制約の強さを調整する。

# 正則化レイヤー - Dropout

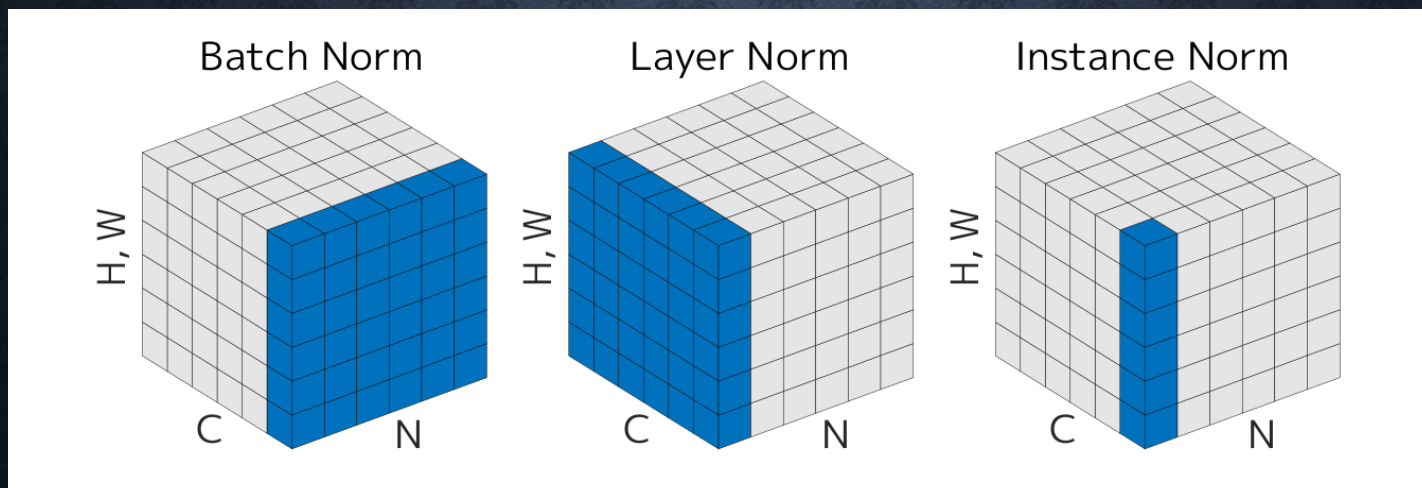
ランダムにノードを削除して学習させる。データ量はそのまま、より疎なモデルで学習することで過学習を抑制。





# 正規化レイヤー

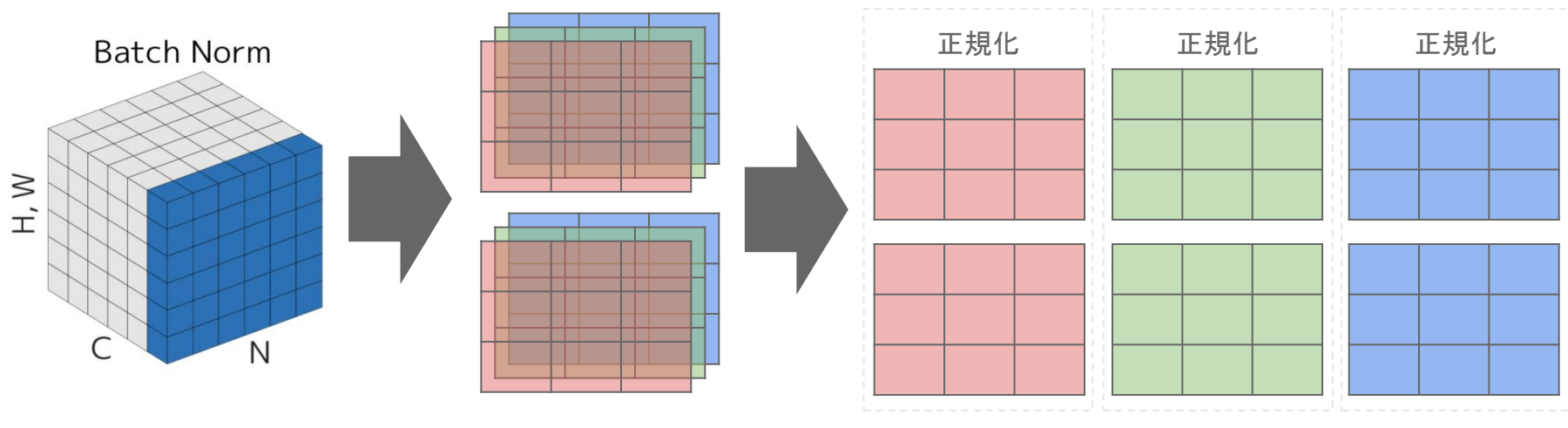
データの正規化(Normalization)。レイヤー間を流れるデータの分布を、何らかの単位で平均が0、分散が1になるように正規化する。



N: ミニバッチ数、C: Channel、H, W: Height / Widthをまとめたもの

# 正規化レイヤー – Batch正規化

$H \times W \times C$ のsampleが $N$ 個あった場合に、 $N$ 個の同一チャンネルが正規化の単位。  
ミニバッチのサイズを大きく取れない場合には、効果が薄くなってしまう。

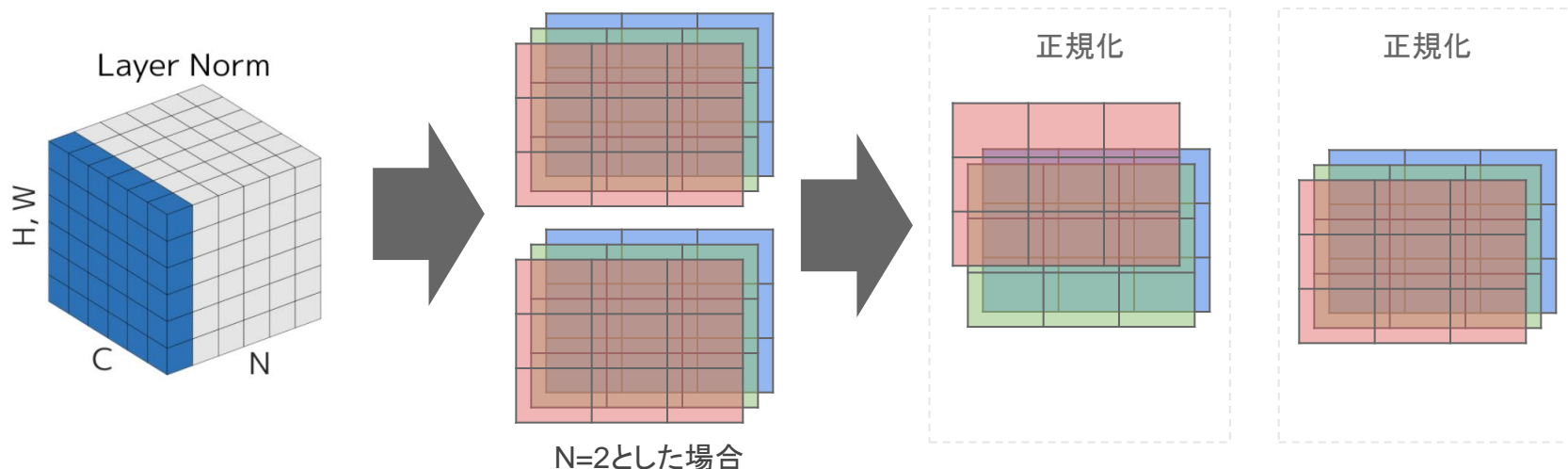




# 正規化レイヤー – Layer正規化

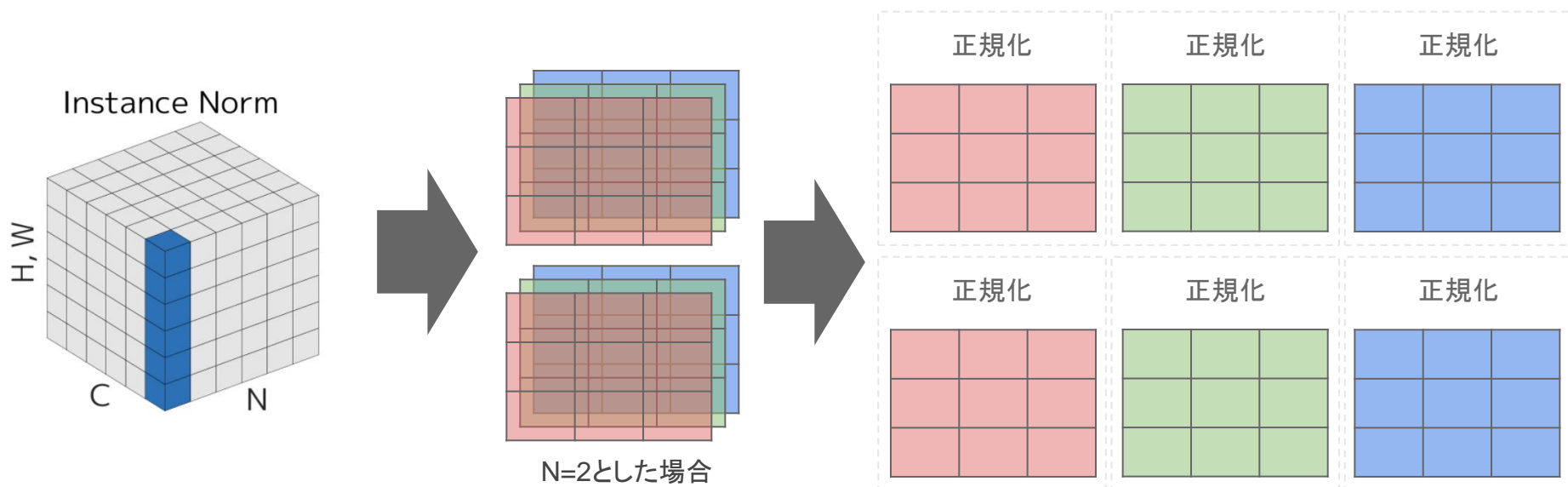
N個のsampleのうち一つに注目。H x W x Cの全てのpixelが正規化の単位。

ミニバッチの数に依存しないので、Batch正規化の問題を解消できていると考えられる。



# 正規化レイヤー – Instance正規化

各サンプルの各チャンネルごとに正規化。  
Batch正規化の際のバッチサイズが1場合と等価。





# ハンズオン

Google Colaboratoryによるハンズオン。

- `regularization.ipynb`
- `layer-normalization.ipynb`
- `dropout.ipynb`