

人工知能・機械学習の工学問題適用 のための課題

和田義孝

近畿大学理工学部機械工学科

破壊力学研究室

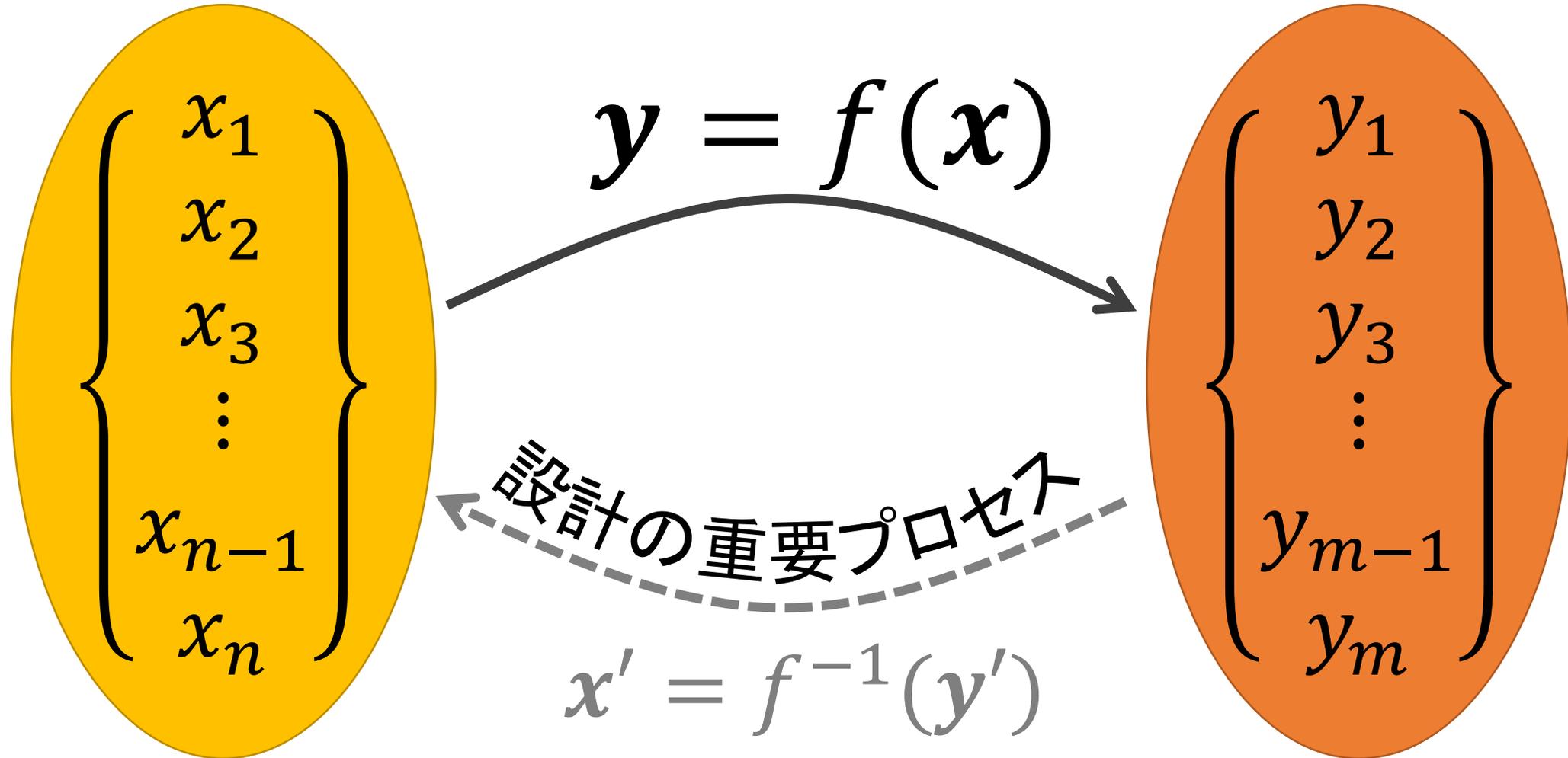


機械学習で設計を援用

- AI-driven CAE, Extended CAE, Real-Time CAE, MLAE(Machine learning Aided Engineering)などと呼ぶのが適切
- 他の分類問題と根本的に異なる点
 - **予測精度高 / 多次元回帰問題に適したネットワーク不明**
- 機械学習（ここではニューラルネットワーク）とは何か
 - 高次元補間関数 $y = f(x)$ ベクトル関数 f が NN
 - 補外はできない? = 巷のうわさ、頭でっかち
- 機械学習の精度は何で決まるのか
 - データの質：何をどれくらい準備：実験計画
 - 絶対に失敗しないCAEによるデータ生成



単なる高次元関数補完 (補外)



一般に $n > m$ なので f^{-1} は求められない

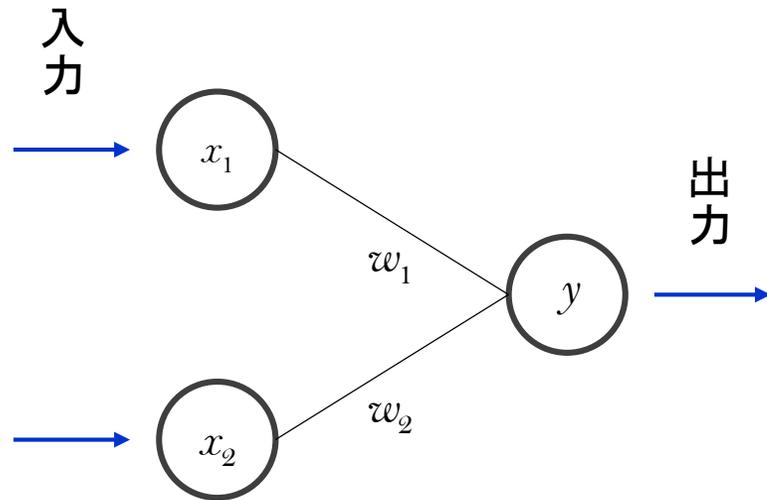


ディープラーニング (深層学習, DL) 1 of 2

ニューラル・ネットワーク(NN)との違い

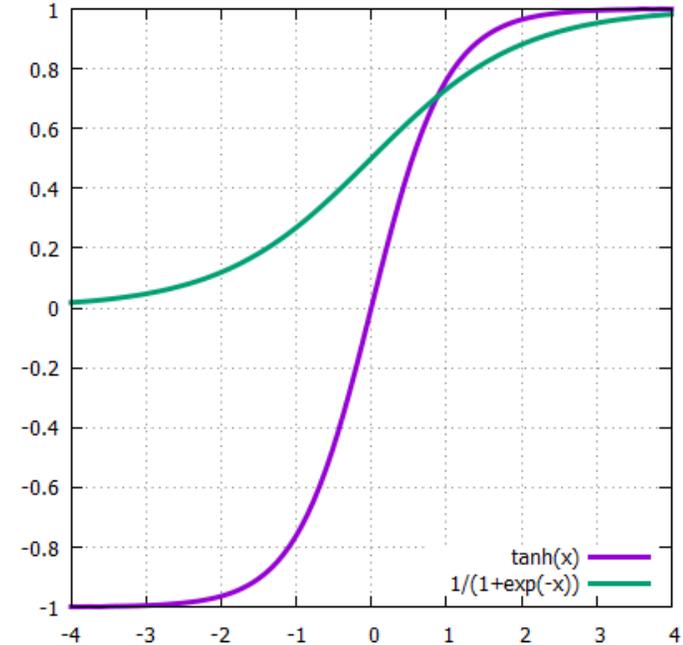
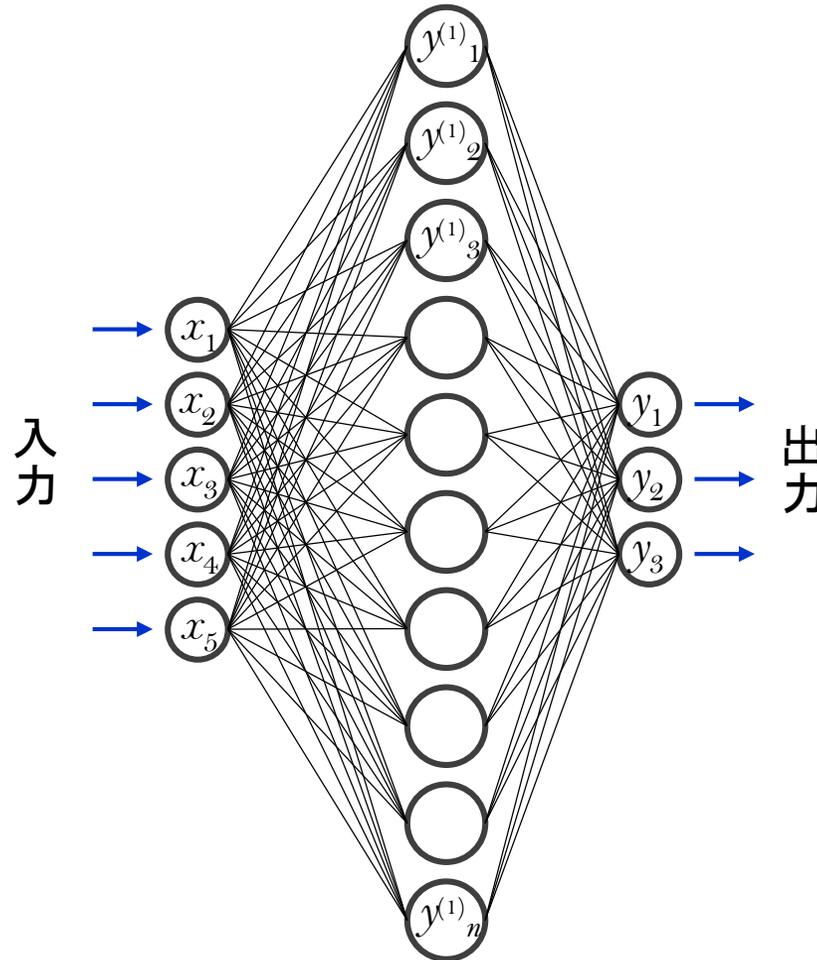
パーセプトロン

ローゼンブラット(1957)



$$y = \begin{cases} 0 & (w_1x_1 + w_2x_2 \leq \theta) \\ 1 & (w_1x_1 + w_2x_2 > \theta) \end{cases}$$

ニューラルネットワーク



シグモイド関数 1

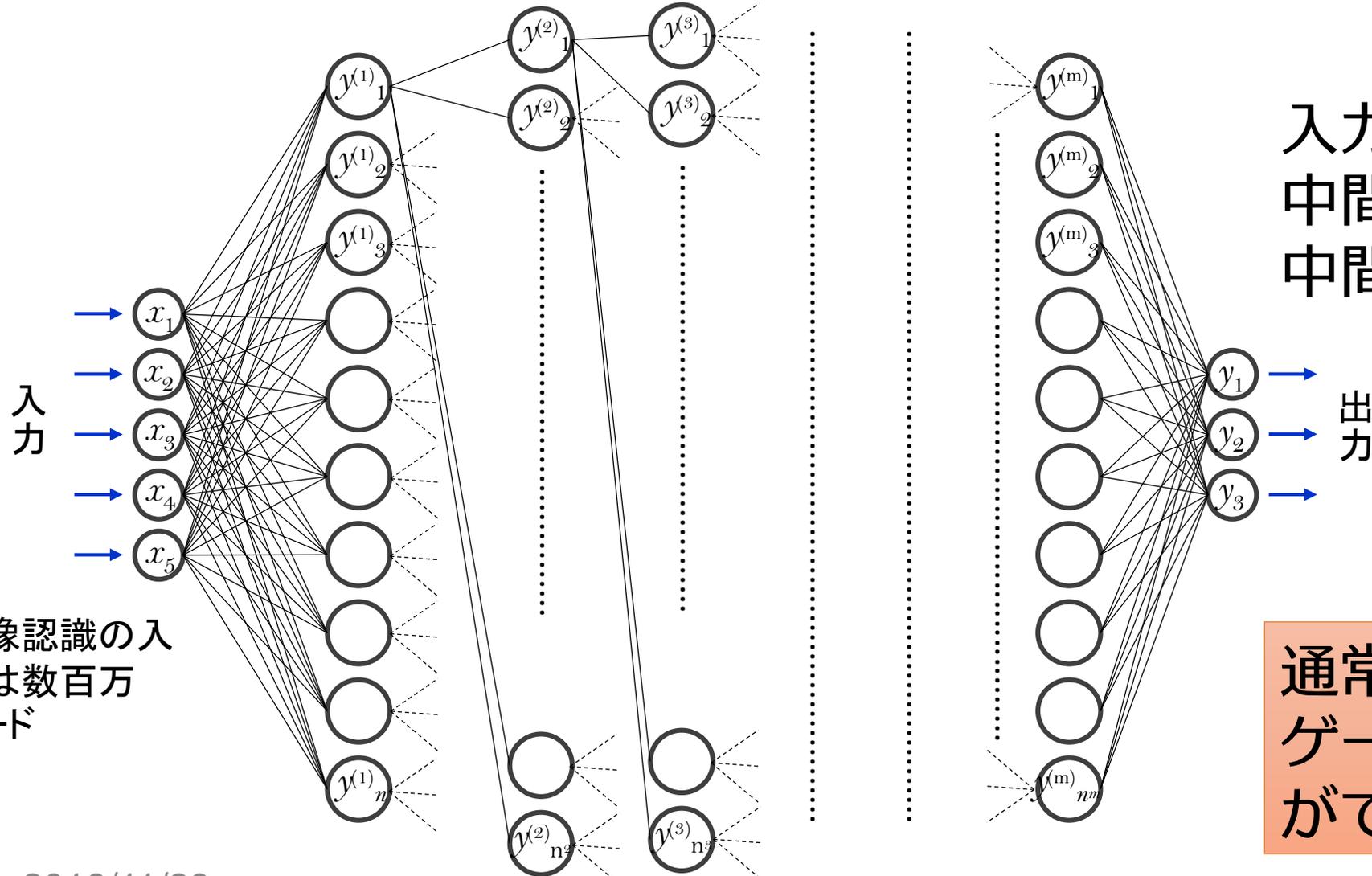
$$h(x) = \frac{1}{1 + e^{-x}}$$

双曲線正接関数

$$h(x) = \tanh(x)$$


ディープラーニング (深層学習, DL) 2 of 2

ニューラル・ネットワーク(NN)との違い



入力ノード数が膨大
中間層の数も膨大
中間層数も数十段

通常のバックプロパゲーション法では学習ができなかった



Expression test: configuration of neural network

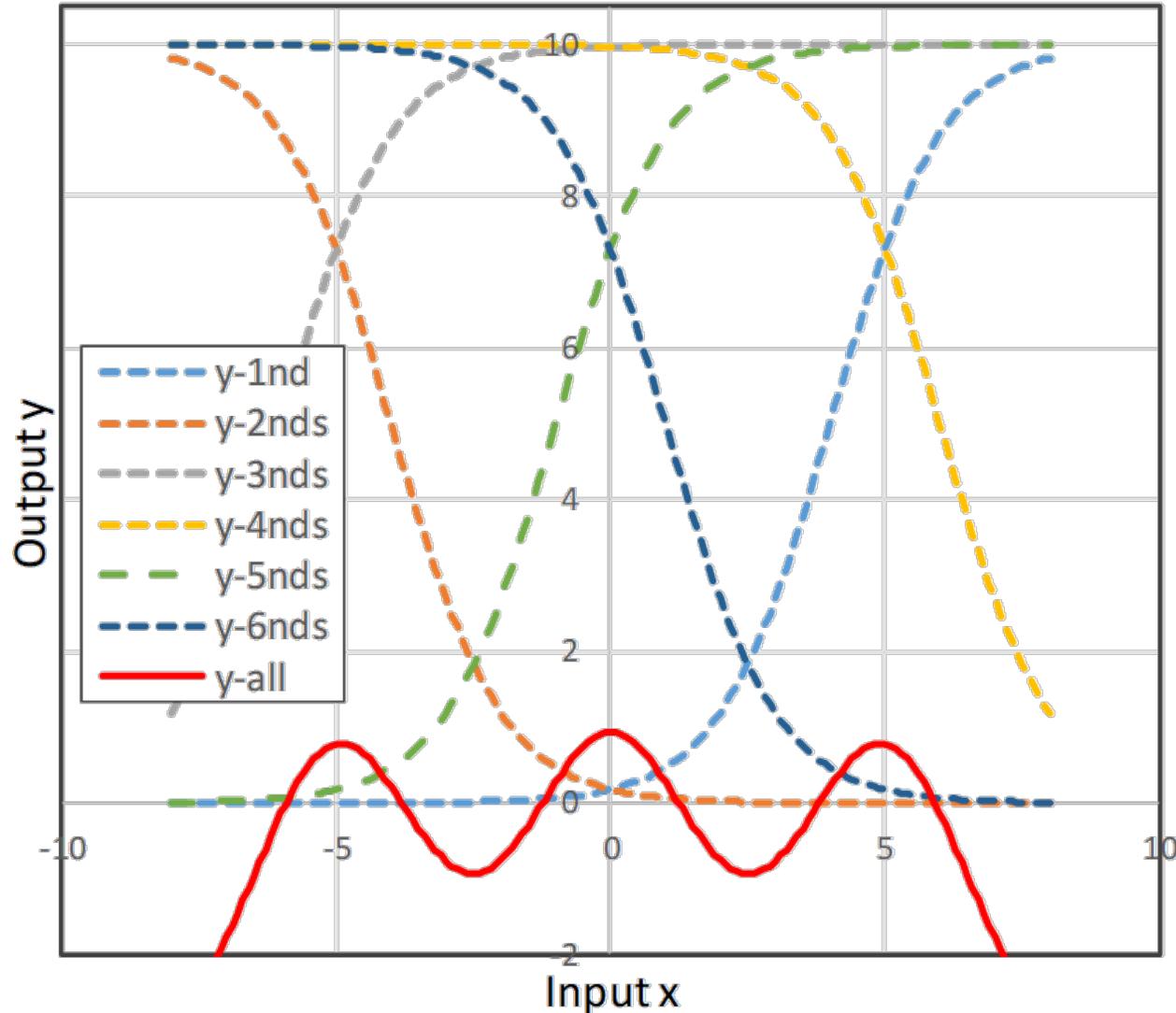
1 input, 1 output, 1 layer, 6 nodes

NN configuration

Input: 1 node
1 layer with 6 nodes
Output: 1 node

Weights and biases

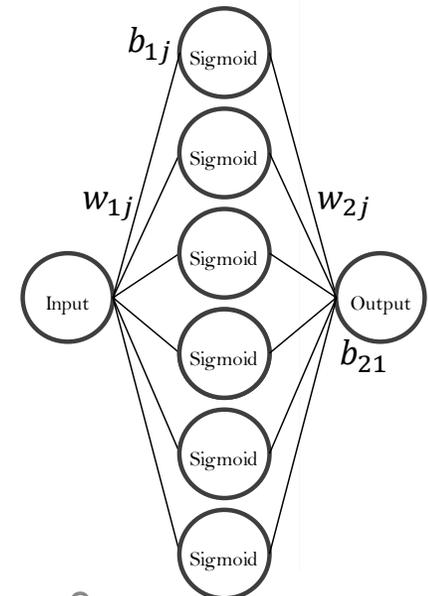
w11	1
w12	-1
w13	1
w14	-1
w15	1
w16	-1
b11	-4
b12	-4
b13	6
b14	6
b15	1
b16	1



Names of weights and biases

Weights and biases

w21	10
w22	10
w23	10
w24	10
w25	10
w26	10
b21	-34

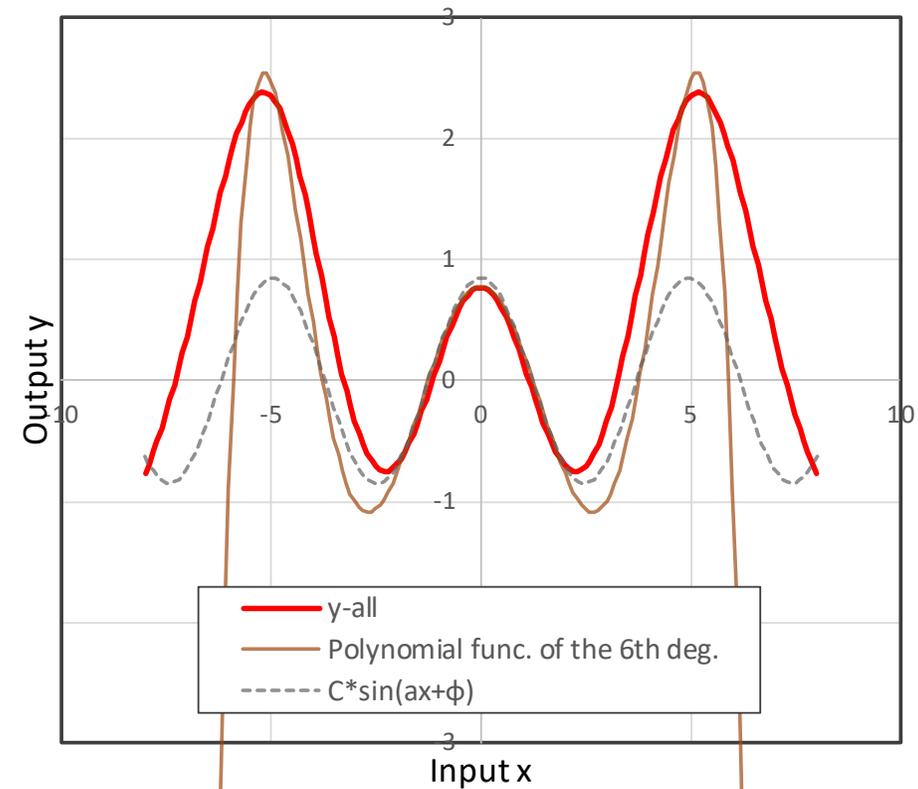
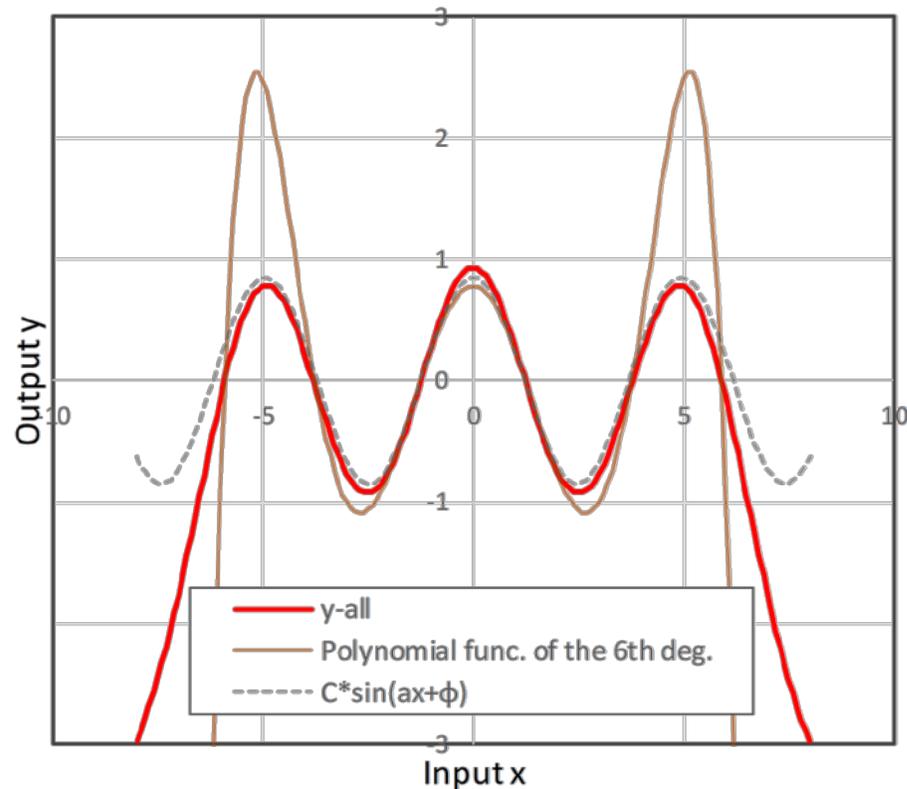


Expression test: configuration of neural network

1 input, 1 output, 1 layer, 6 nodes

Neural network with 1 layer of 6 nodes can express:

- Polynomial function of the 6th degrees,
- 2 cycles of cosine.



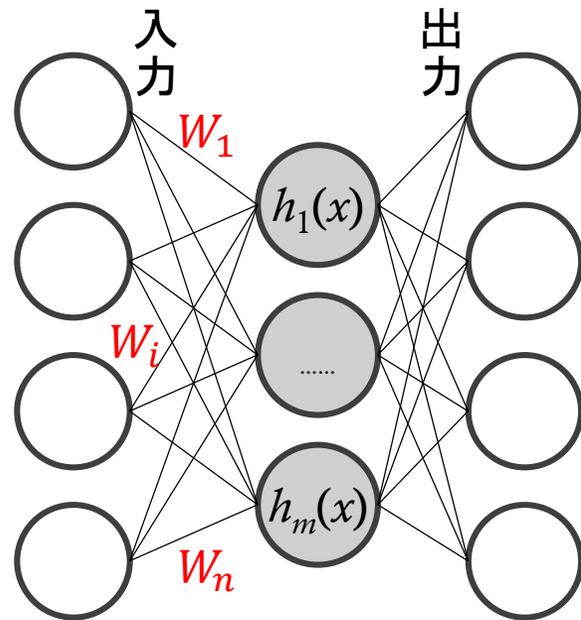
深層学習とは (近畿大和田による定義)

➤ バックプロパゲーション(BP)法の改良および過学習の防止手法も含めた多入力多階層のニューラルネットワーク(NN)の学習手法およびそれらの学習で得たNNの重み係数



バックプロパゲーション (誤差逆伝播学習, BP法)

$$h_j(x) = \frac{1}{1 + e^{-\sum_{i=1}^n (W_i x + b_i)}}$$



1. 学習データの準備

- 入力、出力を1組とし数千から数百万以上のデータセット
- ベクトル量：内積計算のため

2. NN計算し結果ベクトル ←

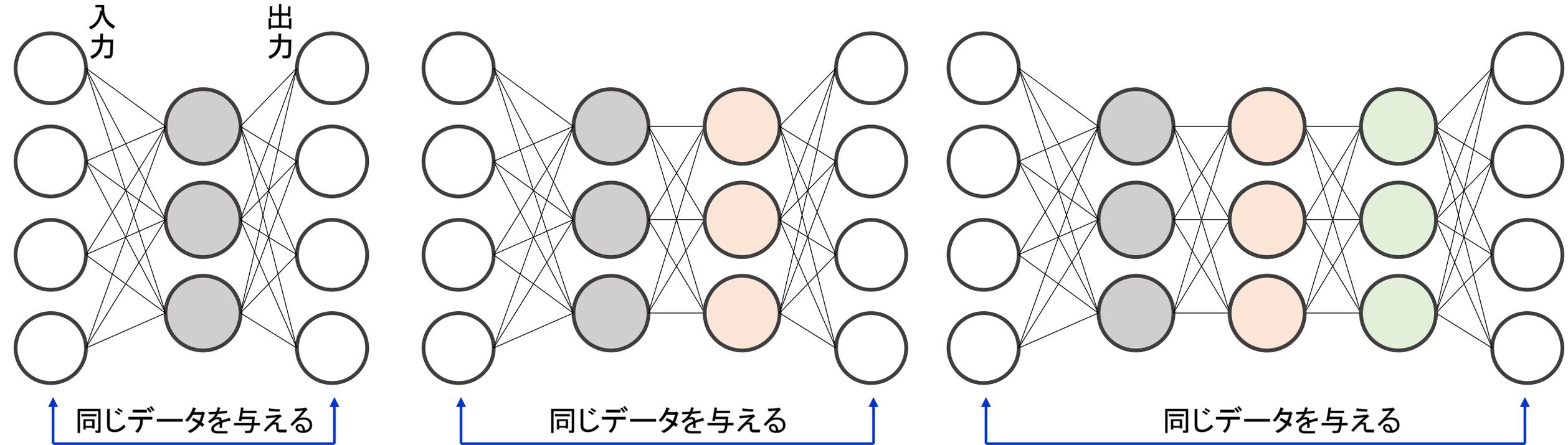
3. 損失関数 (誤差) の勾配

4. 勾配より重み、バイアスの更新

5. 以下2.3.4.を規定回数繰り返す



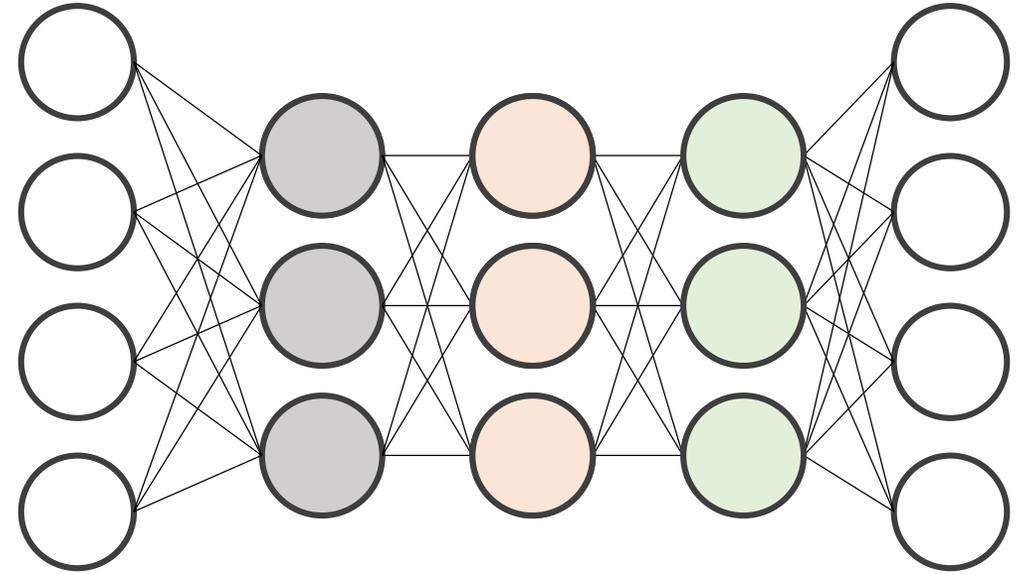
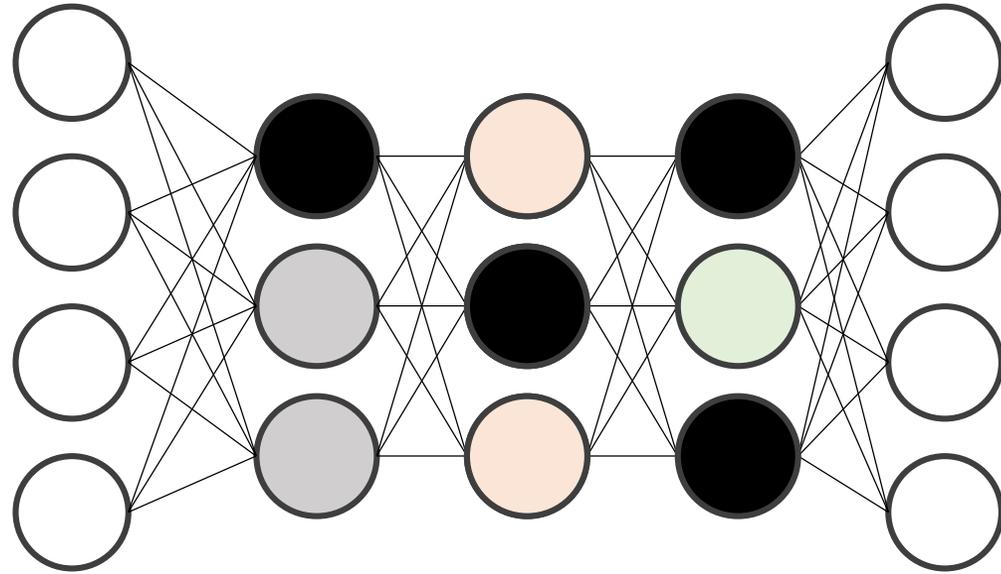
学習方法 1 of 2



- 1層ずつ学習させ多階層においてもBP法可能
- **中間層は入力層より少なくする** (自己符号化、次元圧縮)



学習方法 2 of 2



↑ 誤差データを与える

- 学習時に50%のノードを停止させ学習
- 学習データにノイズ
- 重み、バイアス更新手法：AdaGrad, Adam, Momentum, SGD等

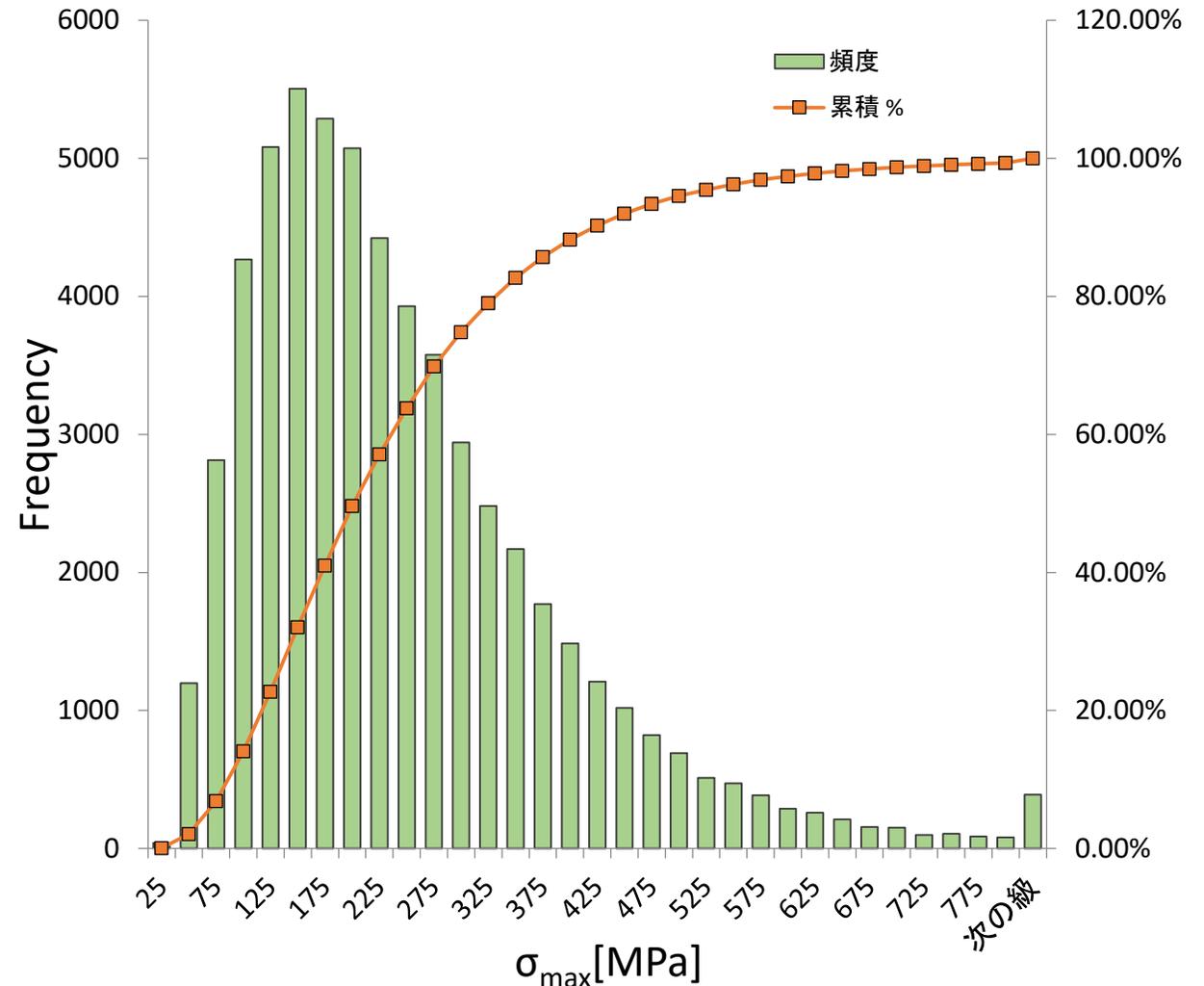


データの質 (最大応力値度数分布)

Imbalanced class: 力学的に適切なデータオーギュメント必要

線形応力解析/評価点における応力	
データ数	約6万
パラメータ数	10
応力最大値	1800 MPa
応力最小値	15 MPa
平均値	235.07MPa
最頻出区間	150-175MPa 5500個

最大応力の度数分布 (25MPa区間)



データの質（組み合わせ 1）

- 期待したいデータは少なくとも2点（2水準）を結ぶようなデータの組み合わせ、できれば3点（3水準）あればよい

例

Name	1	2	3	4	5	6	7	8	9	10	11	12	Eq. Str.
RB05-4411	0.8	100	1	3.7	13.8	10.2	6	0	10	0.5	0.5	32.1	111
RA01-4411	0.8	100	1	3.7	13.8	10.2	6	0	10	0.5	0.5	64	53
RA03-4411	0.8	100	1	3.7	13.8	10.2	6	0	10	0.5	0.5	94	32

- パラメータ⑫が3水準あり理想的、このような組み合わせが大多数であることが望ましい、**C3と呼称**する



データの質（組み合わせ2）

- 2水準のC2で972、3水準のC3で18しかない
- 1 3次元のデータに対して組み合わせの頻度がこれだけ偏っていると**学習可能かどうか判別不可**
- 計画的なデータ準備が不可欠
 - 実験計画の知見の有効活用、L9 (3^4), L27 (3^{13}) 直交表など

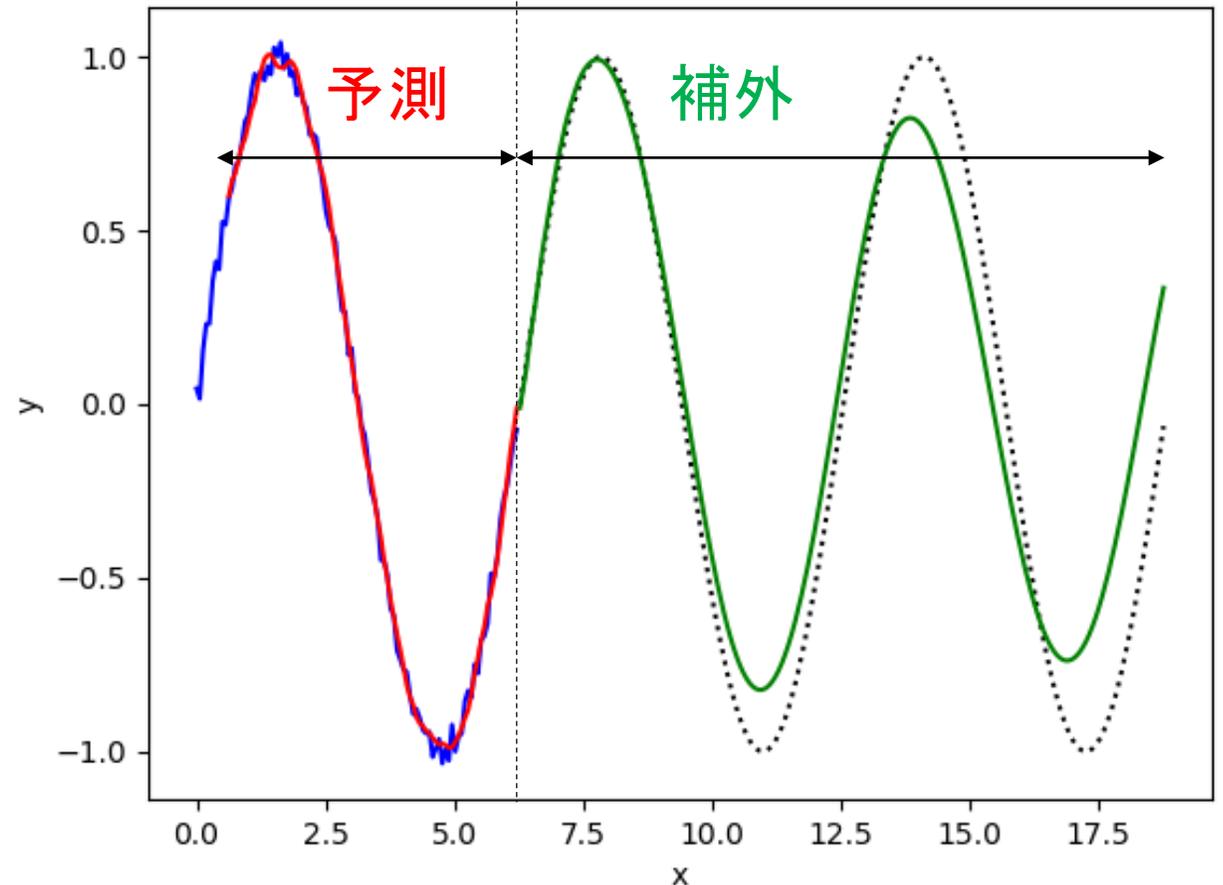
分類	頻度 (全データ数58986)
C1	57994
C2	974
C3	18
C4以上	0



基本問題：補外できるか？

- 学習データ外 ($2\pi \sim 6\pi$)
予想
- 補外範囲は未学習
- 学習データ ($0 \sim 2\pi$)
- 予測 ($\sim 2\pi$)

Hidden Ly	Batch Size	Lr. Rate	Epochs
300	300	0.001	100



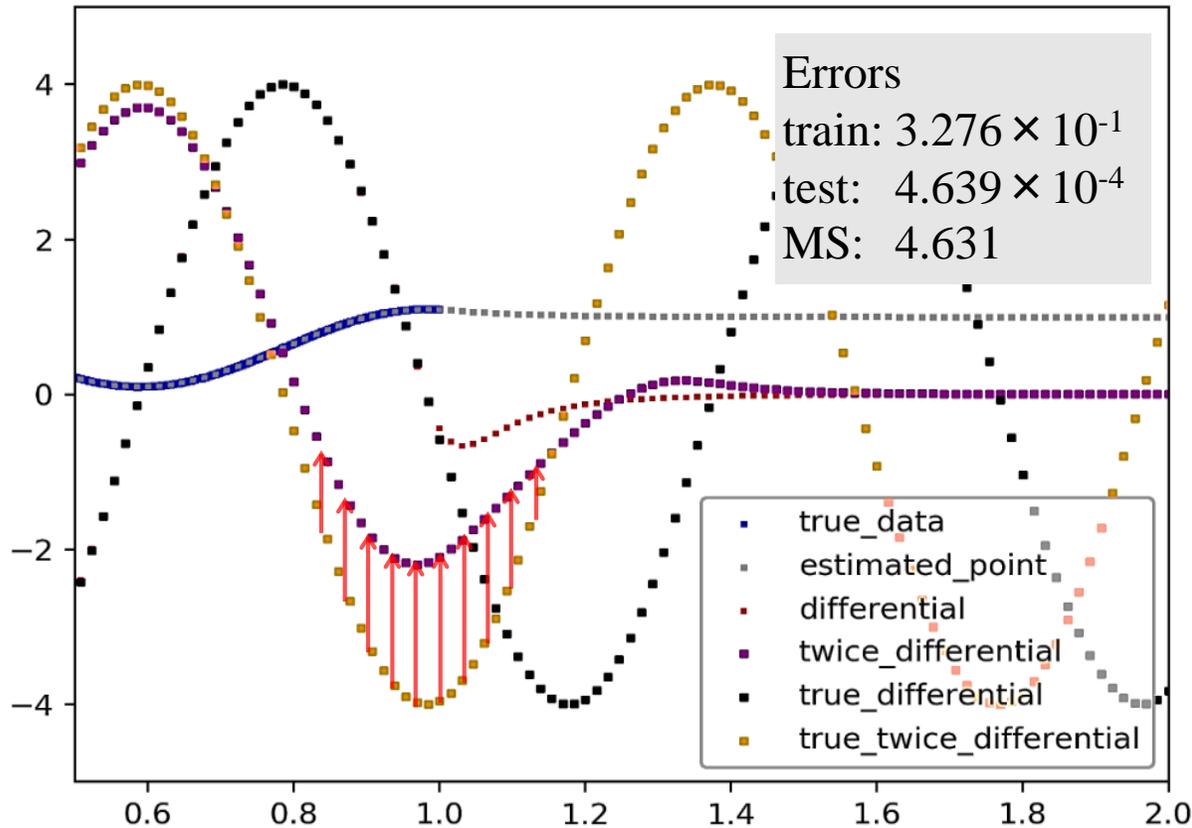
$y = \sin(x)$ の予測と補外



Result of CGC 1 of 2

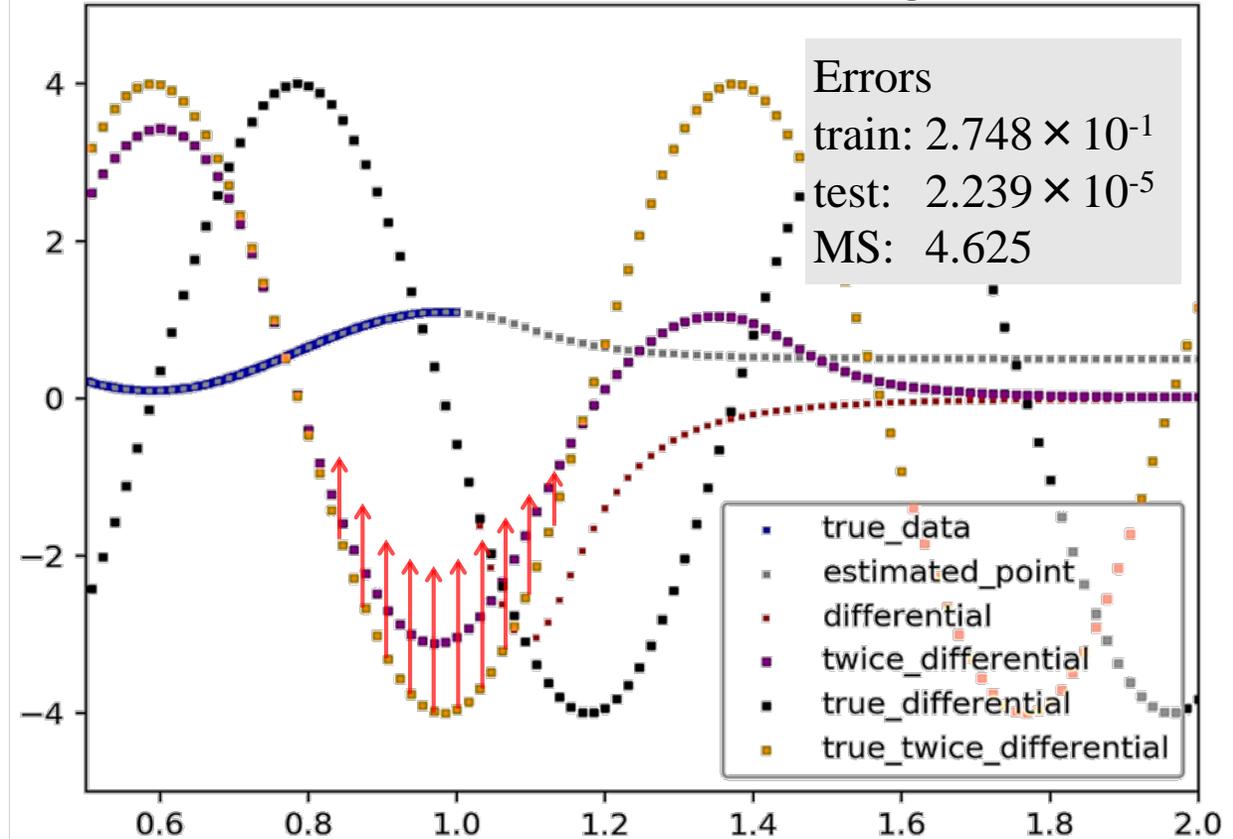
Hidden layer: 9(3), nodes on layer 11(3), 2 million trainings

NN (2 million trainings): 16h



Insufficient for extrapolation

CGCNN (2 million trainings): 8h



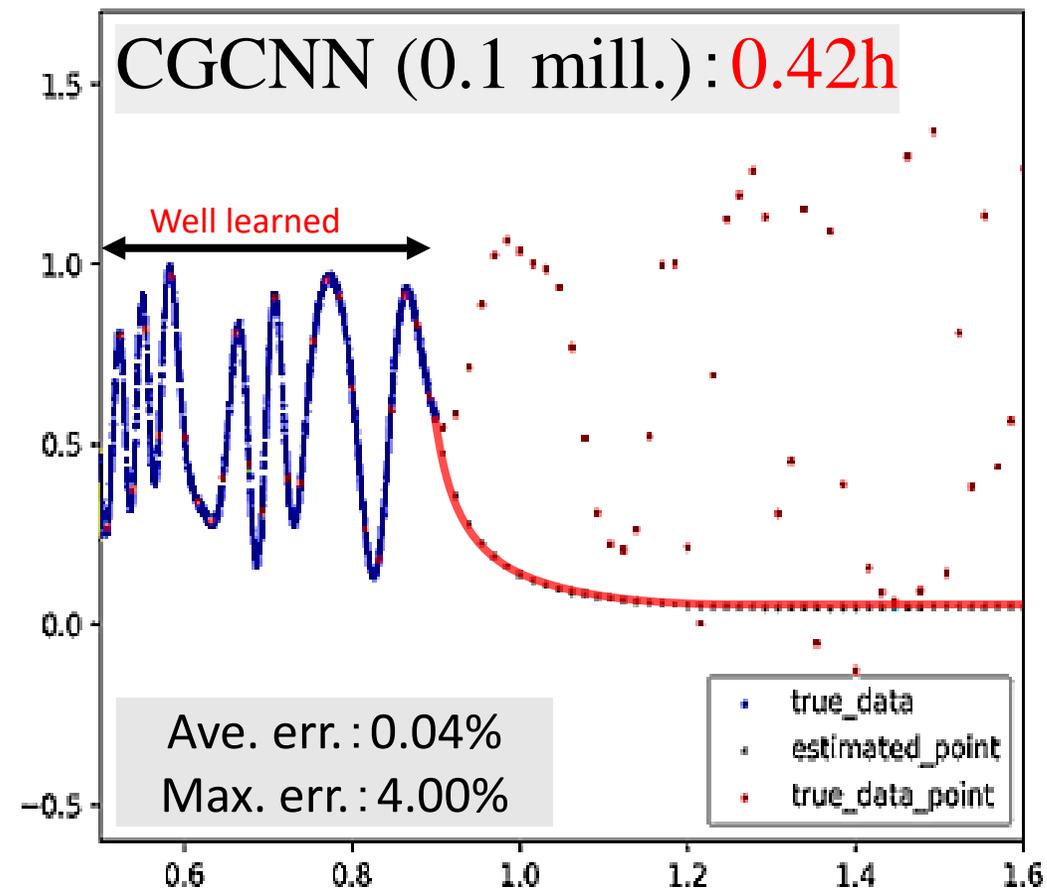
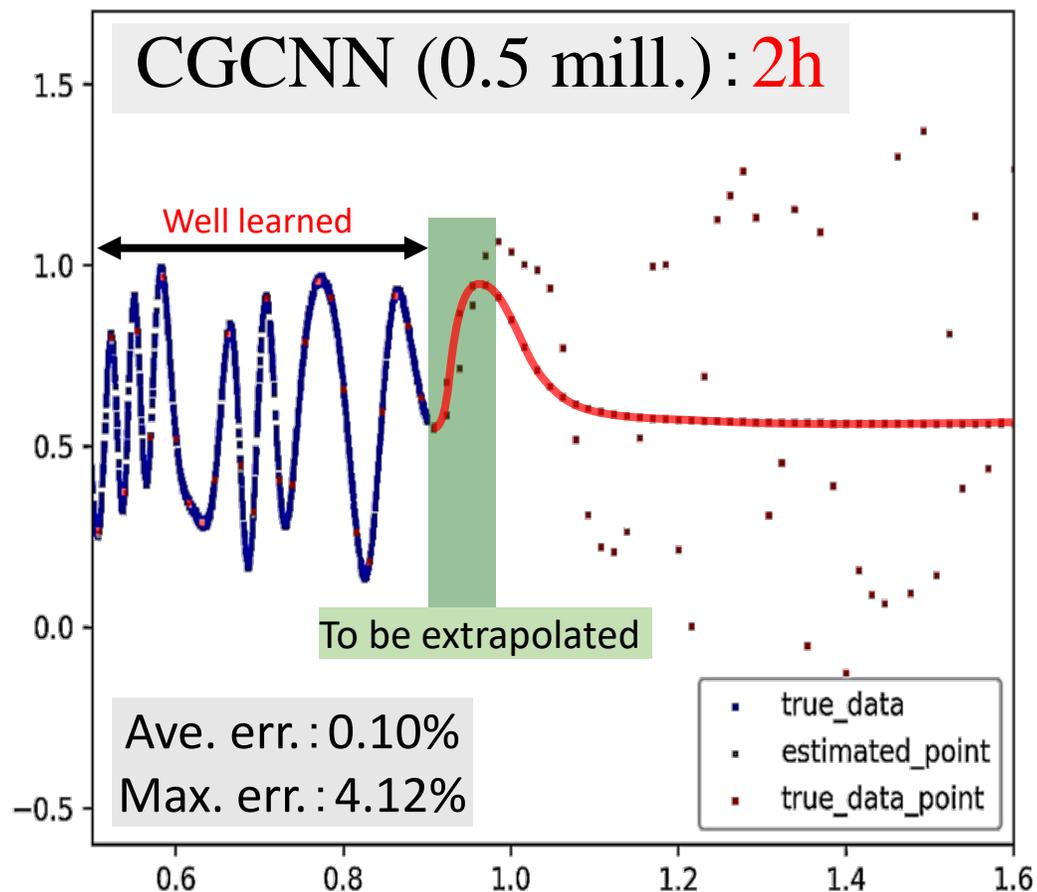
Better result



Result of CGC 2 of 2

Hidden layer: 9(5), nodes on layer: 9(5), 0.5 million trainings

- Teaching function : $f(x) = \sin\left(\frac{35x}{10}\right) + \cos\{25(x-1)\cos(20x)\}\sin(x/2) + 0.6$



予測の質保証のために(HQCへ向けて1)

	数値シミュレーション	機械学習/深層学習
対象物理現象≡ データ分布・特徴	支配方程式 (微分方程式、経験的)	データに依存 (データ統計分析)*
方程式の求解≡学 習手法	連立一次方程式の解法 (ソルバー)	学習方法、最適化関数、活 性化関数などの組合*
離散化(メッシュ) ≡ネットワーク構成 表現能力	物理量変化に対応する離 散化モデル(メッシュ)	ニューラルネットワーク構 成(MLP,CNN,LSTM)*
その他手法依存	境界条件処理、MPCなど	転移学習など [△]



予測の質保証のために(HQCへ向けて2)

機械学習/深層学習のデータに必要な条件

データそのもの
に対する要件

表したい現象を十分に含んでいるか
(頻度が高い・低い情報があるなどの特徴について
理解し対策が分かっているか)

学習させるた
めの要件

現状の学習方法に対して十分なデータ拡張
されているか

物理現象
の記述は
データに依
存

データ・現象・
機械学習への
理解

過剰な精度を要求していないか
(物理現象への理解があるか≠学習数少・ば
らつきがあるなどに関わらず)

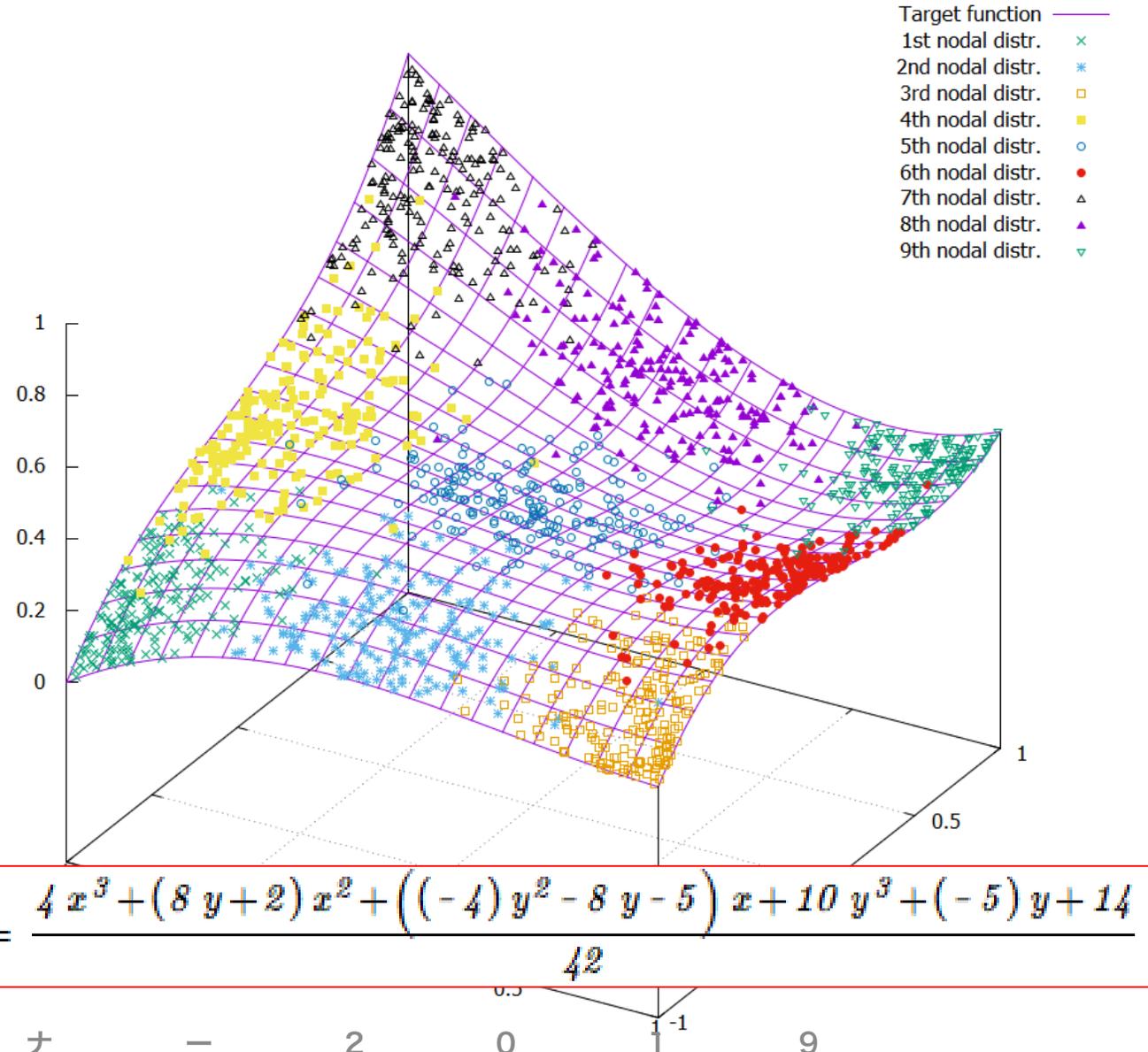
十分条件はあ
るか?

よくわからない
少なくとも共通手順が定まらないとHQCへ向かえない

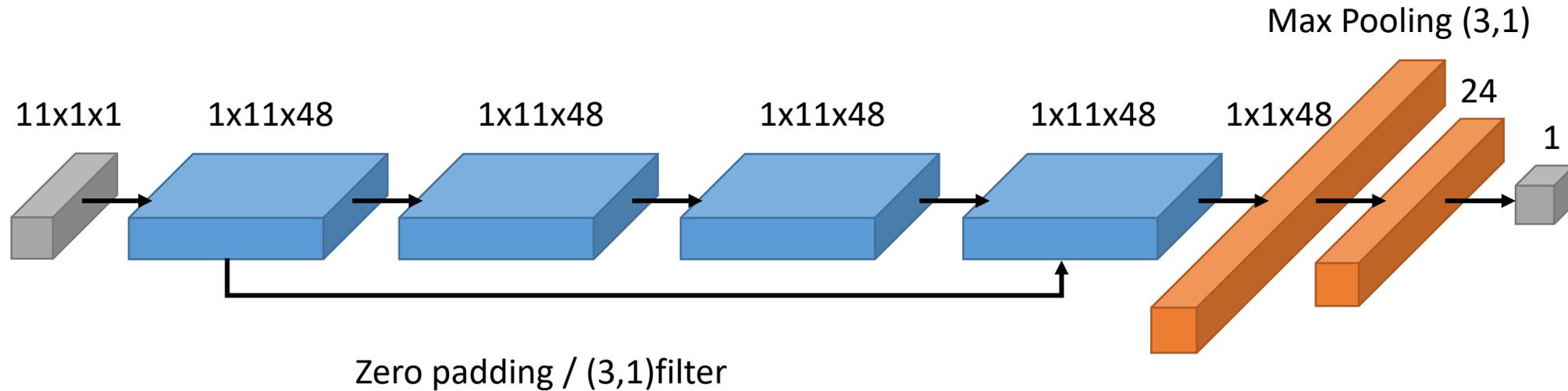


基本問題

- 図のような多項式で表される関数を予測
- 何を目的とするか
 - 学習に適切なパラメータの数を調査
 - 最低限の入力パラメータはx, y
 - 最大は多項式の項数（項数は8、性能評価のためダミー項として11入力も追加）
 - この範囲において極値が1か所存在する
 - $(x=0.549, y=0.561, z=0.240)$
 - データ密度、データ数により学習の可否を検討



Convolutional NN with Residual Network: vector input



- Activation function: hyperbolic tangent
- Optimizer: Adam with learning rate = 10^{-5}
- Residual Network: 1st convolution layer to 4th convolution layer
- Batch normalization: applied
- Drop out: applied at 20% rate



組み合わせの考え方

- 必要最小限のパラメータがあれば原則学習は行われるが膨大なデータが必要
- 学習に必要なデータが少なくとも学習を加速させるための手段
 - 関係のあるパラメータのオーギュメント（パラメータ積や微分なども有効な場合あり）
 - 入力側にノイズの付与
 - 物理法則に必要なニューラルネットワークの構成

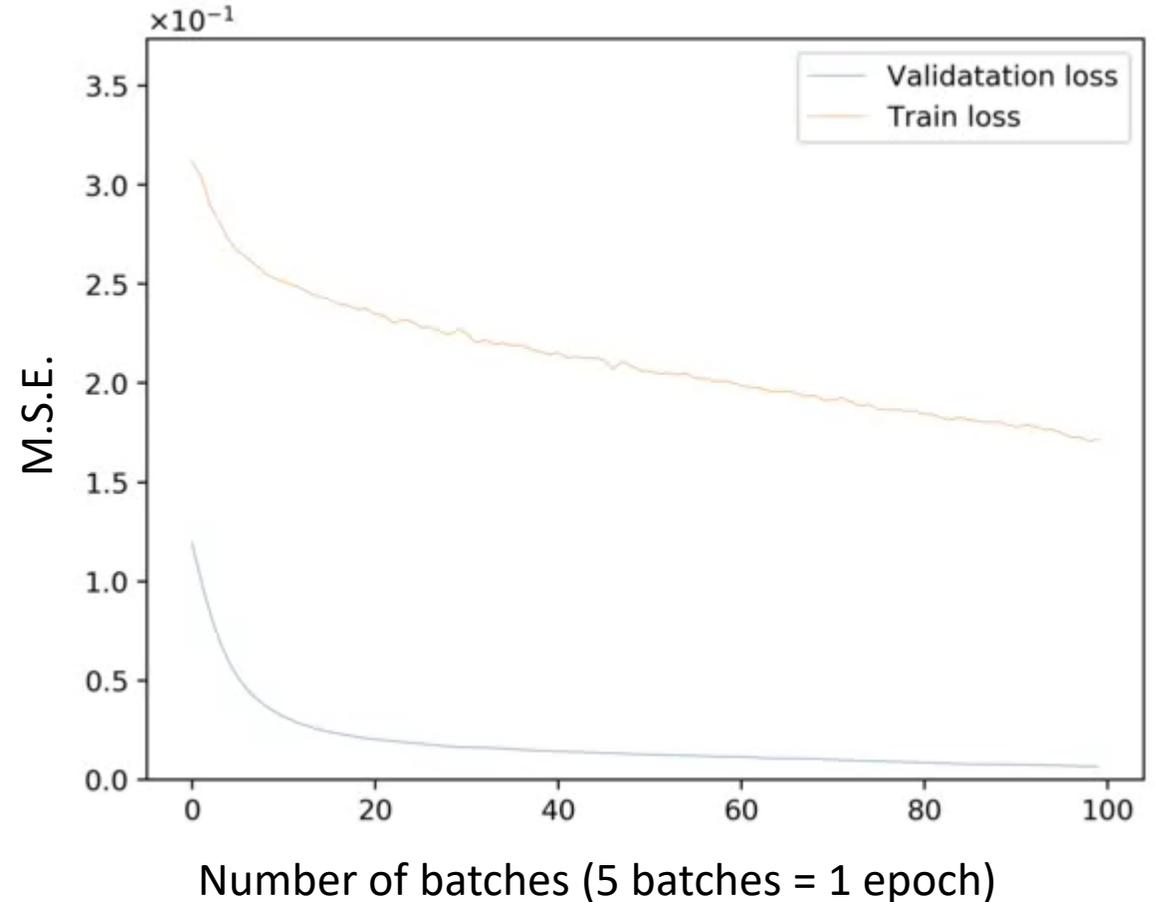
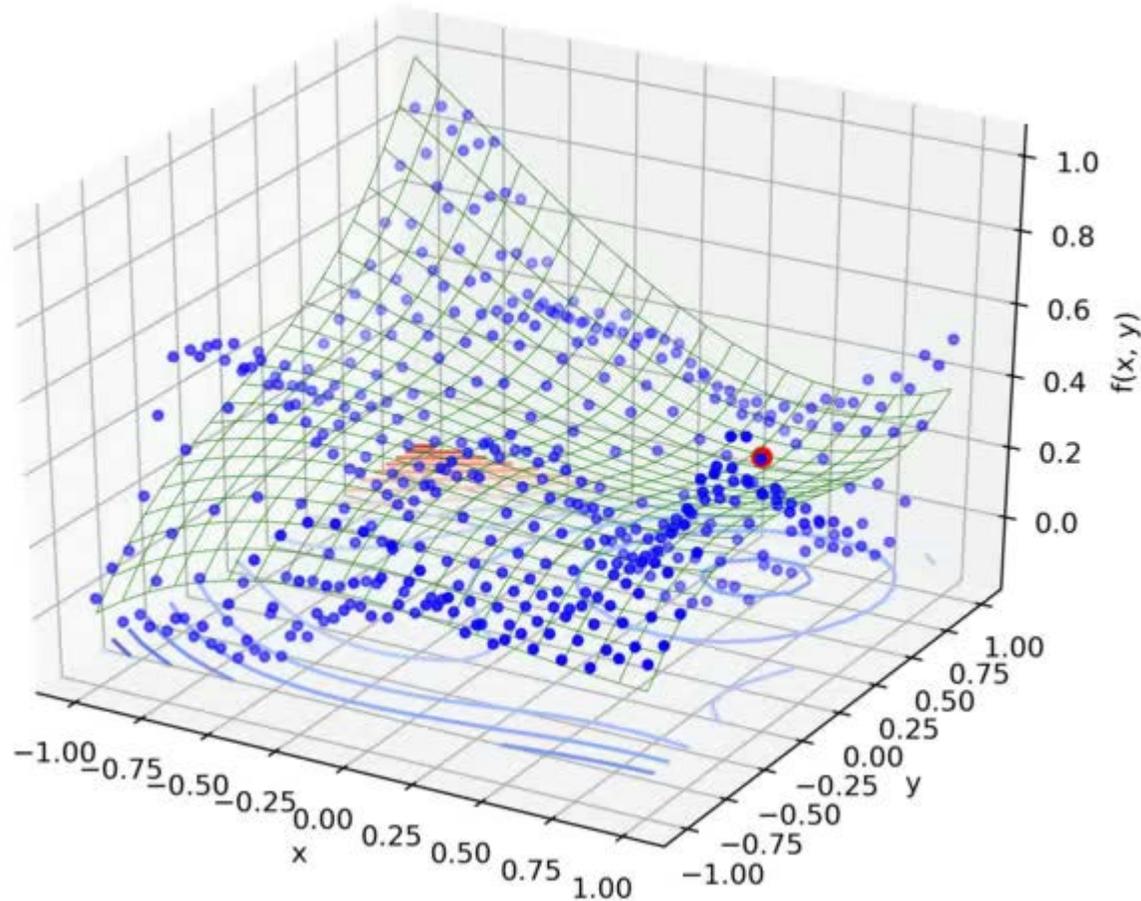
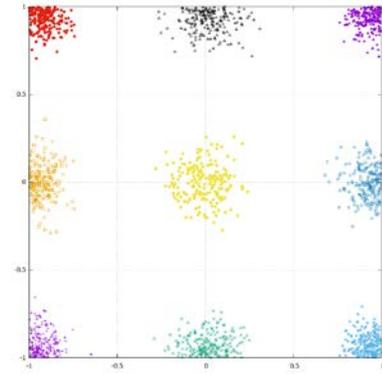
元のデータはxとyのみで関数は記述可能であるが、多項式の同定のためには膨大なニューロンが必要となるため可能性のある組み合わせは先に計算で求める

	1	x	x ²	x ³	x ⁴
1	1	x	x ²	x ³	x ⁴
y	y	xy	x ² y	<u>x³y</u>	x ⁴ y
y ²	<u>y²</u>	xy ²	<u>x²y²</u>	x ³ y ²	x ⁴ y ²
y ³	<u>y³</u>	xy ³	x ² y ³	x ³ y ³	x ⁴ y ³
y ⁴	y ⁴	xy ⁴	x ² y ⁴	x ³ y ⁴	x ⁴ y ⁴

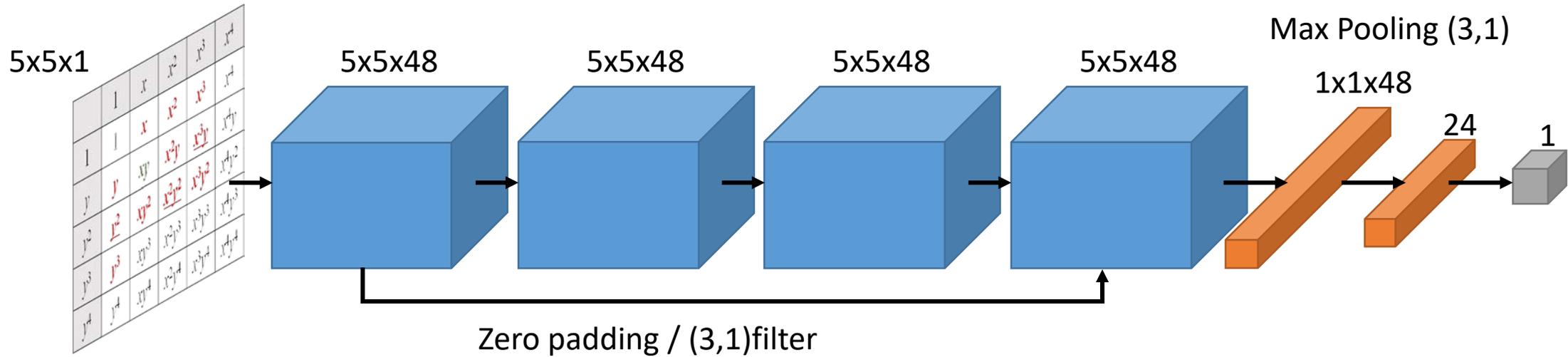
$$f(x, y) := \frac{4x^3 + (8y + 2)x^2 + ((-4)y^2 - 8y - 5)x + 10y^3 + (-5)y + 14}{42}$$



$\sigma=0.1, n=200, n_t=n \times 9=1800$
Data-augmented : total data sets is 90,000



Convolutional NN with Residual Network: image-based(matrix) input

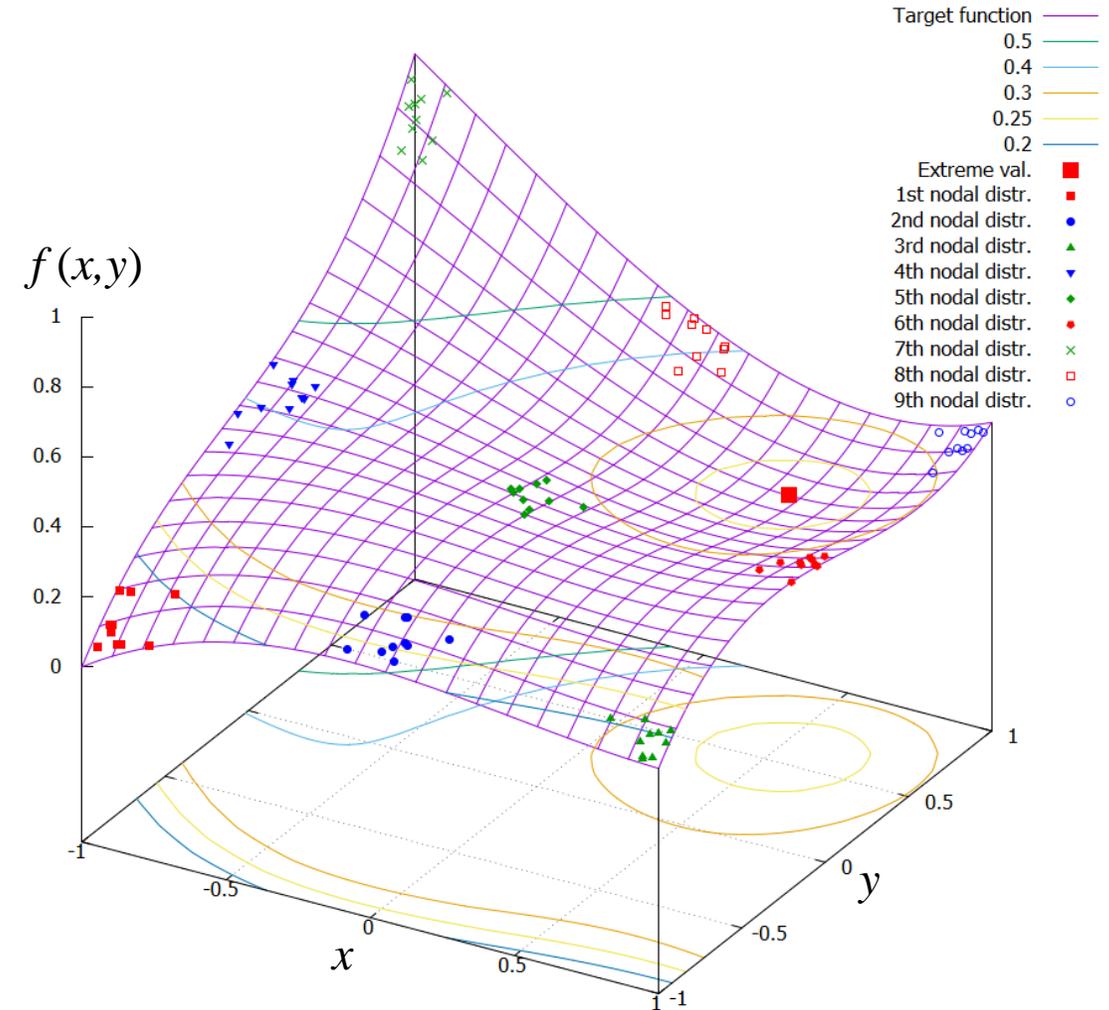
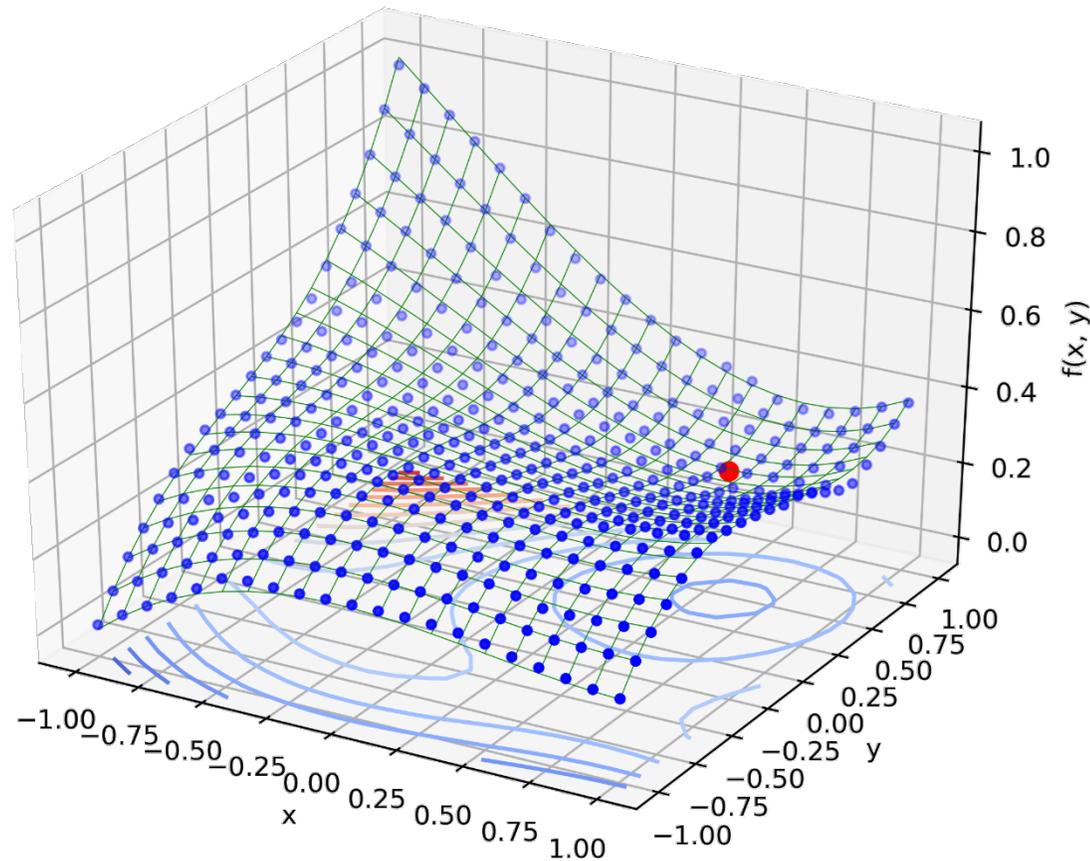


- Activation function: hyperbolic tangent
- Optimizer: Adam with learning rate = 10^{-4}
- Residual Network: 1st convolution layer to 4th convolution layer
- Batch normalization: applied
- Drop out: applied at 20% rate



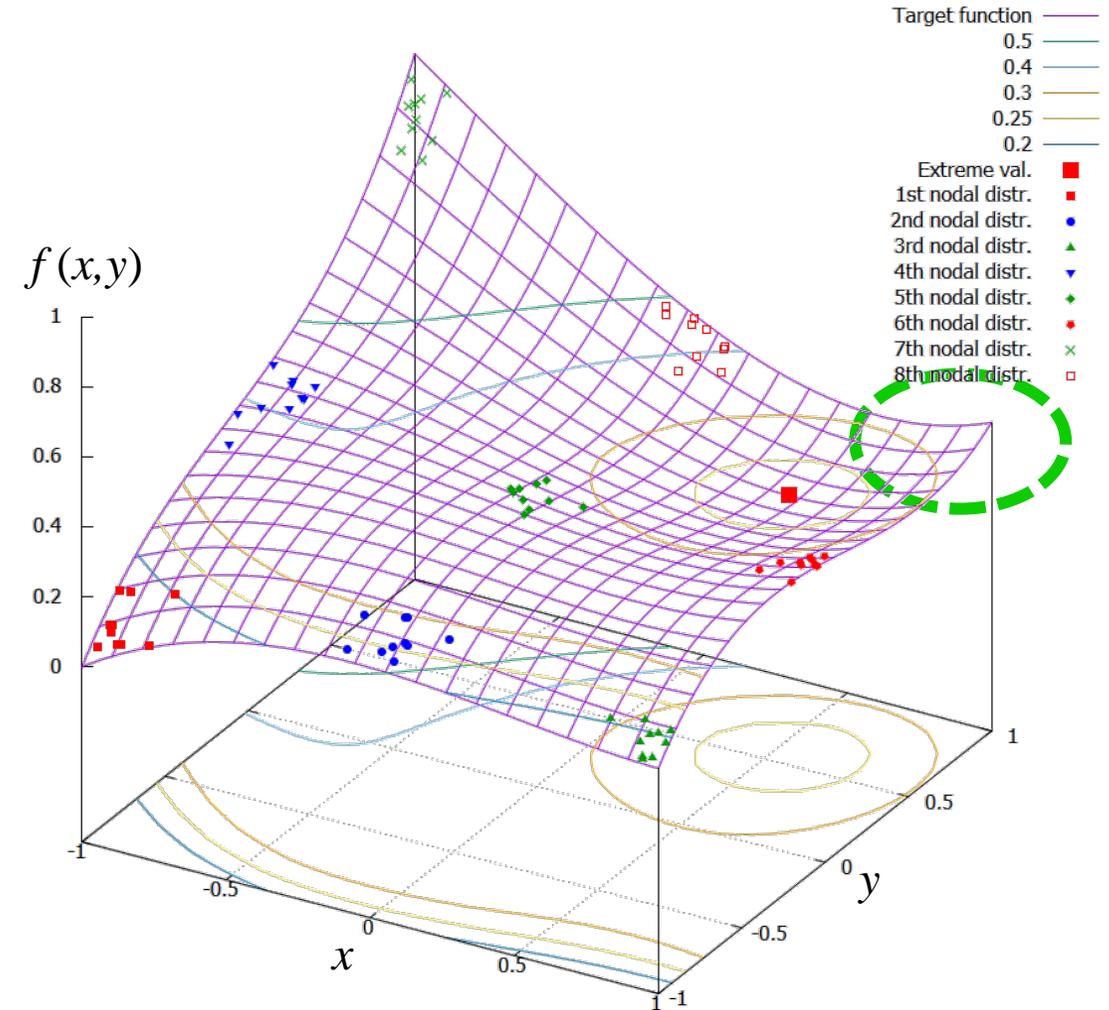
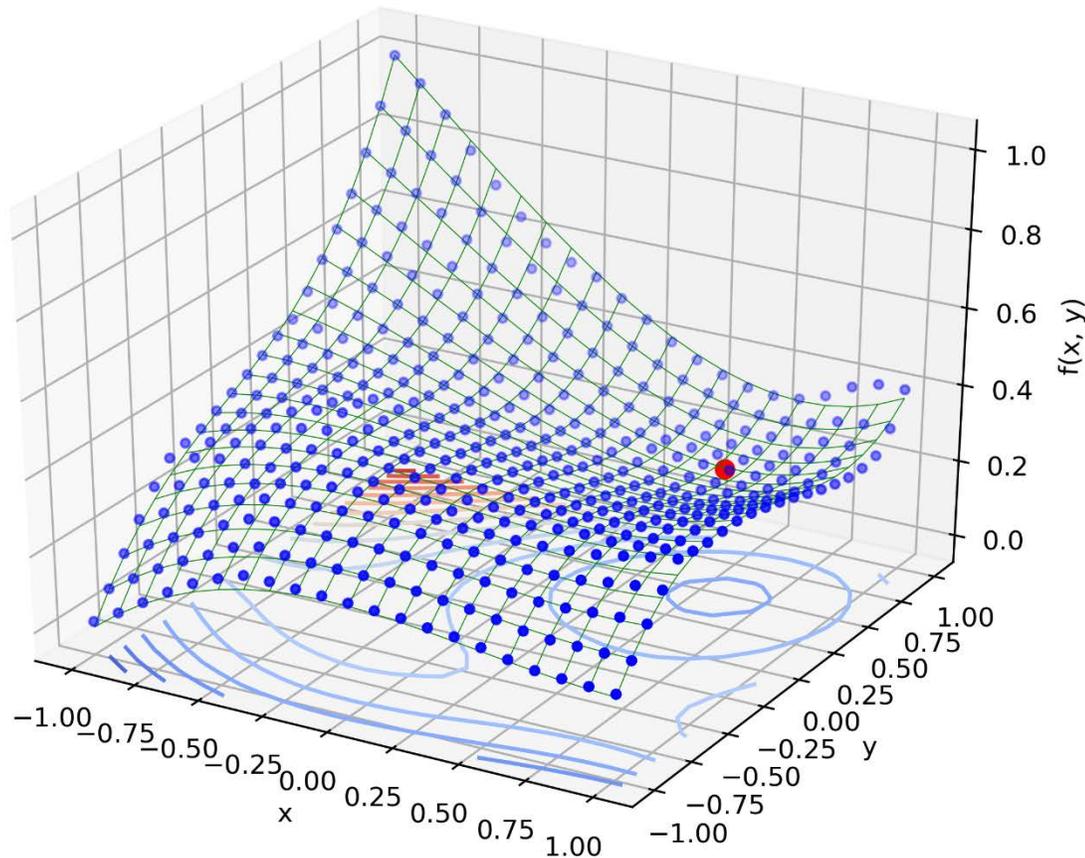
$\sigma=0.1, n=10, n_t=n \times 9=90$

Data-augmented : total data sets is 90,000



$\sigma=0.1, n=10, n_t=n \times 8=80$

Data-augmented : total data sets is 80,000



重要な知見

- BP法が大域最適化の考え方 => **データの性質に依存大**
- データの一様性 / Uniformity of training data
 - 一様乱数 \neq 一様期待量、**パラメータ空間を一様**にしても出力値（学習対象）は**正規分布**や**ワイブル分布**
 - 部分要因計画（直交表）、ラテン超方格法（ $L_9(3^4)$, $L_{27}(3^{13})$ ）などの**相関を考慮したデータ分布**
- 特徴の付与（与えるデータの特徴を考慮）
- データの水増し / Data augmentation
 - ノイズの付加
 - 過学習防止 / 汎化性能向上



Neural Networkを用いた き裂進展予測

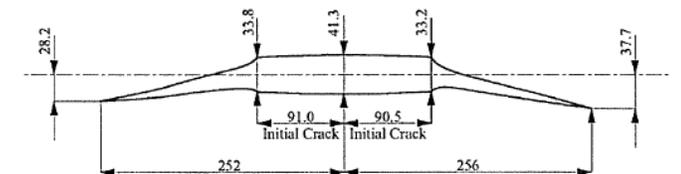
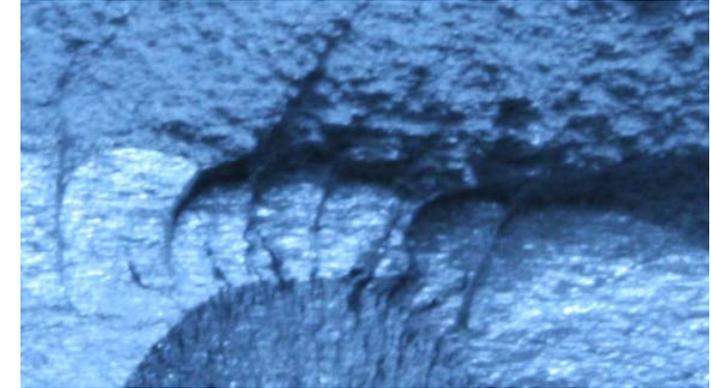


Classification of numerical simulation for crack propagation

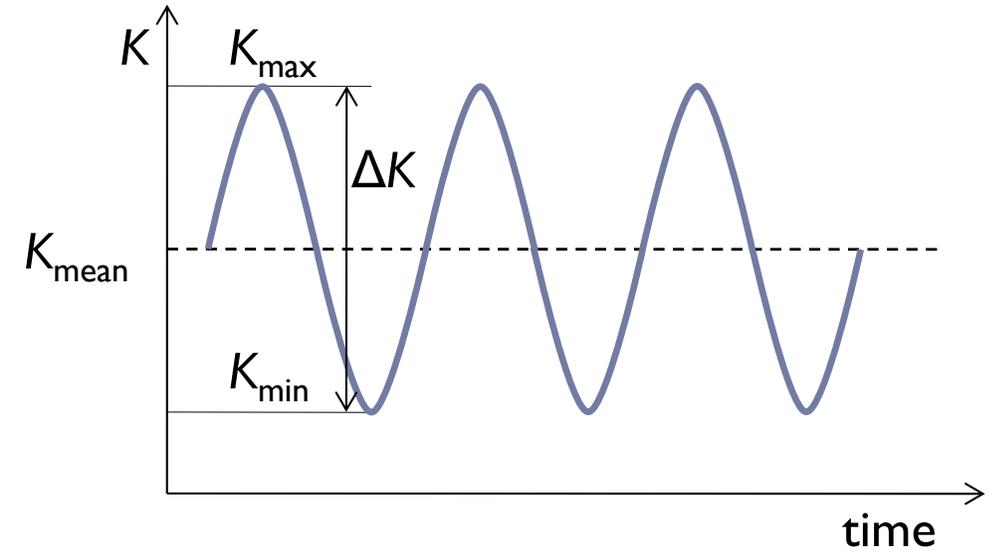
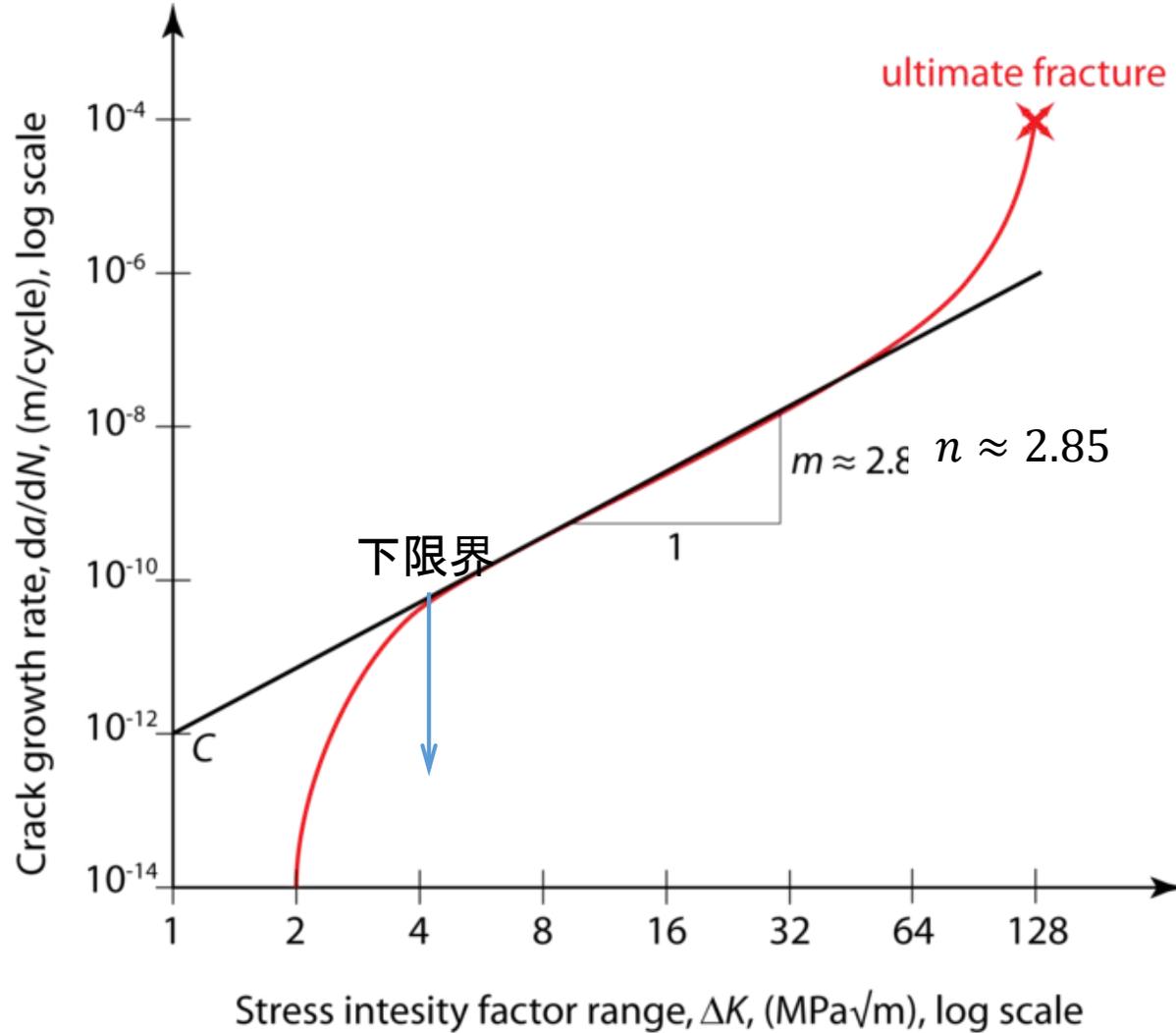
K: stress intensity factor, K_c : brittle fracture toughness

J: J-integral, J-R curve: ductile fracture toughness

Static or dynamic	Type of fracture	Yielding zone size	Loading / number of cycles	Difficulties for numerical simulation	Parameters for fracture prediction
static	brittle	small	monotonic	-	K_c
			cyclic / high	multiaxial loadin, no accurate direction criteria in 3-d	K_c , Paris' law, MTS criterion, etc.
	ductile	Large	monotonic	very large deformation, no direction prediction	J_c
			cyclic / low	ΔJ for crack propagation addition to the above	J_c , ΔJ , J-R curve, etc.
dynamic	Similar to brittle		assumed monotonic	similar to brittle	dynamic J_c , $K_{II}=0$ criterion



Paris-Erdogan則



$$\frac{da}{dN} = C (\Delta K_{eq})^m$$



出典: <http://en.wikipedia.org/wiki/File:ParisLaw.png>

2019/11/29

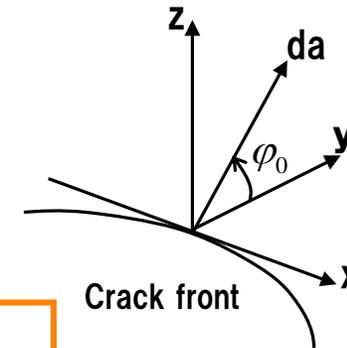
き裂進展則

Crack growth direction: Richard's criterion

$$\varphi_0 = \mp \left[A \frac{|K_{II}|}{K_I + |K_{II}| + |K_{III}|} + B \left(\frac{|K_{II}|}{K_I + |K_{II}| + |K_{III}|} \right)^2 \right] \quad (*1)$$

$$\varphi_0 < 0^\circ \text{ for } K_{II} > 0 \text{ and } \varphi_0 > 0^\circ \text{ for } K_{II} < 0 \text{ and } K_I \geq 0$$

$$A=140^\circ, B=-70^\circ.$$



Crack growth rate: Paris's law

$$da / dN = C (\Delta K_{eq})^n \quad (*2)$$

$$K_{eq} = \frac{K_I}{2} + \frac{1}{2} \sqrt{K_I^2 + 4(\alpha_1 K_{II})^2 + 4(\alpha_2 K_{III})^2} \quad (*1)$$

$$\alpha_1 = K_{ic} / K_{IIc} \text{ and } \alpha_2 = K_{ic} / K_{IIIc}$$

$$\alpha_1 = 1.155, \alpha_2 = 1.0$$

Material

A533B steel $E=206\text{GPa}$, $\nu=0.3$
 $C=1.67 \times 10^{-12}$, $n=3.23$

*1 H. A. Richard, M. Fulland, M. Sander S. N. (2005): Blackwell Publishing Ltd. Fatigue Fract Engng Mater Struct, Vol. 28, pp. 3-12.

*2 P.C.Paris and F.Erdogan, J. bas. Engng. m Trans. ASME, Ser. D, 85, 1963, 528-533

疲労き裂進展に関する法則および物理量

破壊力学パラメータ

- 応力拡大係数 K_I, K_{II}, K_{III} , 破壊靱性値：
 K_{IC}

パリス則：き裂進展量とサイクルの関係

- 材料定数： C, m $\frac{da}{dN} = C(\Delta K_{eq})^m$

き裂の形状(ベクトル, 多直線)： \mathbf{p}_i

FEMで計算される諸量

- 変位ベクトル： \mathbf{u}
- 応力テンソル： $\boldsymbol{\varepsilon}$
- ひずみテンソル： $\boldsymbol{\sigma}$

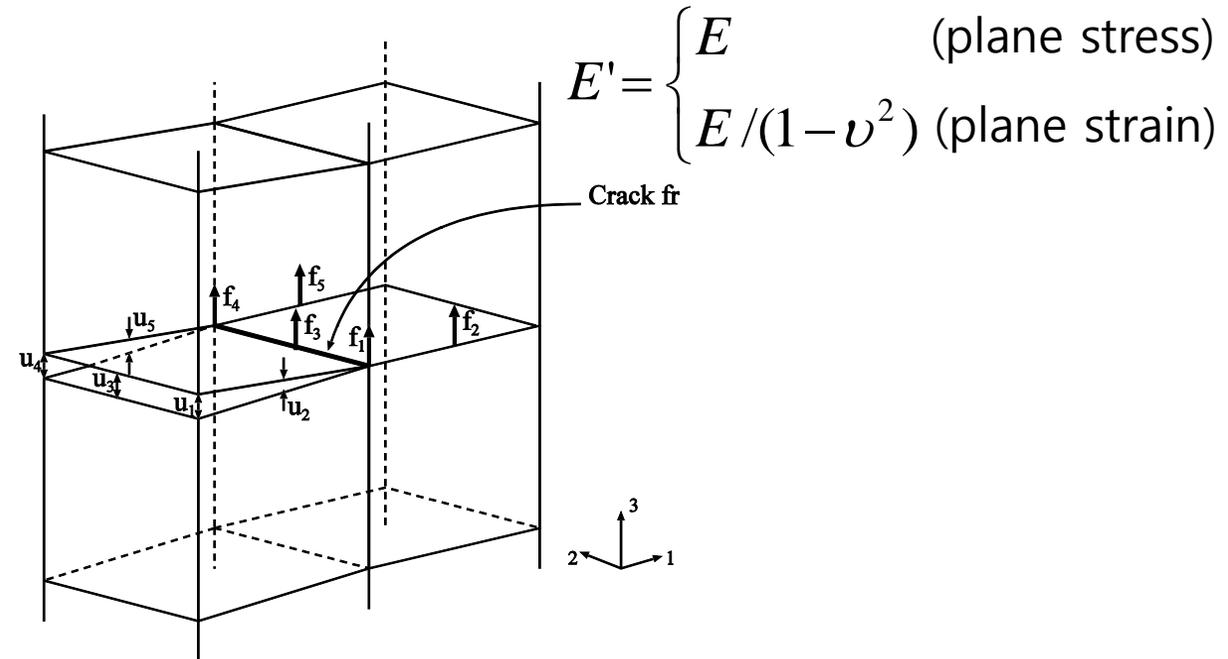
材料物性

- ヤング率： E , ポアソン比： ν

応力拡大係数の計算方法

$$G_I = \frac{1}{2A} \sum_{i=1}^5 u_i^3 f_i^3 \quad K_I = \sqrt{G_I E'}$$

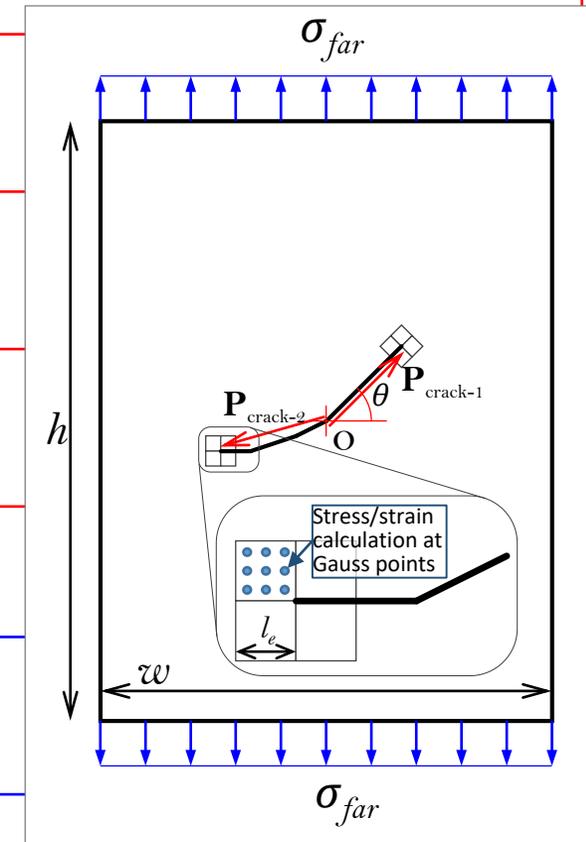
$$G_{II} = \frac{1}{2A} \sum_{i=1}^5 u_i^1 f_i^1 \quad K_{II} = \sqrt{G_{II} E'}$$



Training data

Input

Common	σ_{far} da	Stress MPa Crack prop. dist. m	\mathbf{p}_{crack}	Crack tip position (Origin at crack center, vector)
Phase1	K_I, K_{II}	S.I.F. MPa/mm ^{1/2}		
Phase2	$\sigma_x, \sigma_y, \tau_{xy}$	Stress MPa		
Phase3	<u>Only crack tip position is input data</u>			



Output

All phases	da	Crack prop. vector, dN	cycles per a cal. step
------------	----	------------------------	------------------------



Training conditions

Phase 1, 2 and 3

Number of training sets

: 4884

(Initial crack angle $\theta=0,15,30,45,60$ deg.)

✘ $\theta=45$ up to 75 steps, 60 up to 30 steps

Pre-training : 25 times (dropout : 10%)

Fine-training : 20,000

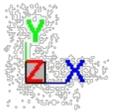
Hidden layers : 6 (every layer has 60 nodes)

Verification model

Initial crack angle 30°



Results from initial to 30th step



s-version FEM

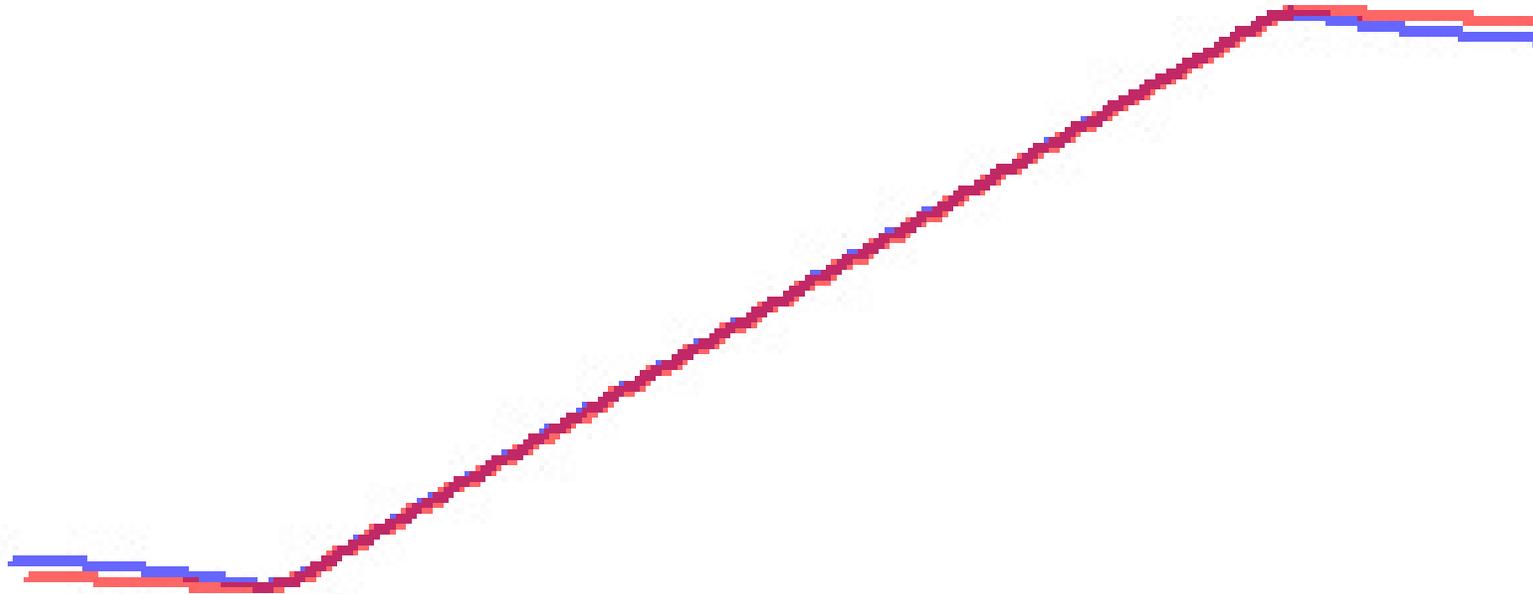
Deep Learning

2019/11/29

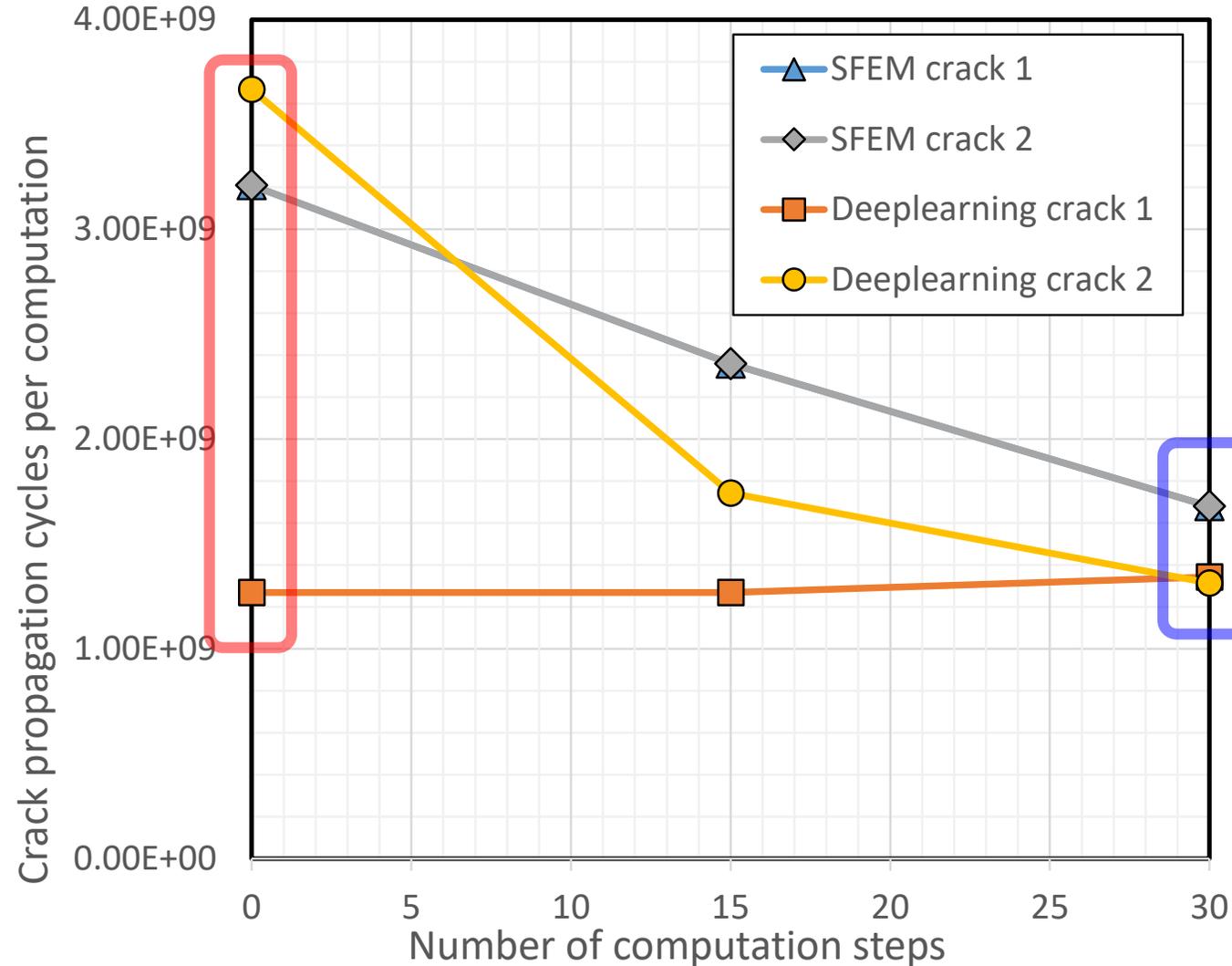
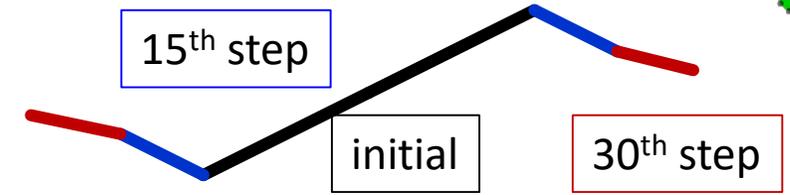
Comparison at 30th step on phase 1

Blue: correct crack shape by s-FEM

Red: predicted crack shape by deep neural network



Results of number of cycles $\Delta a=0.15\text{mm}$



Converged to the same cycles

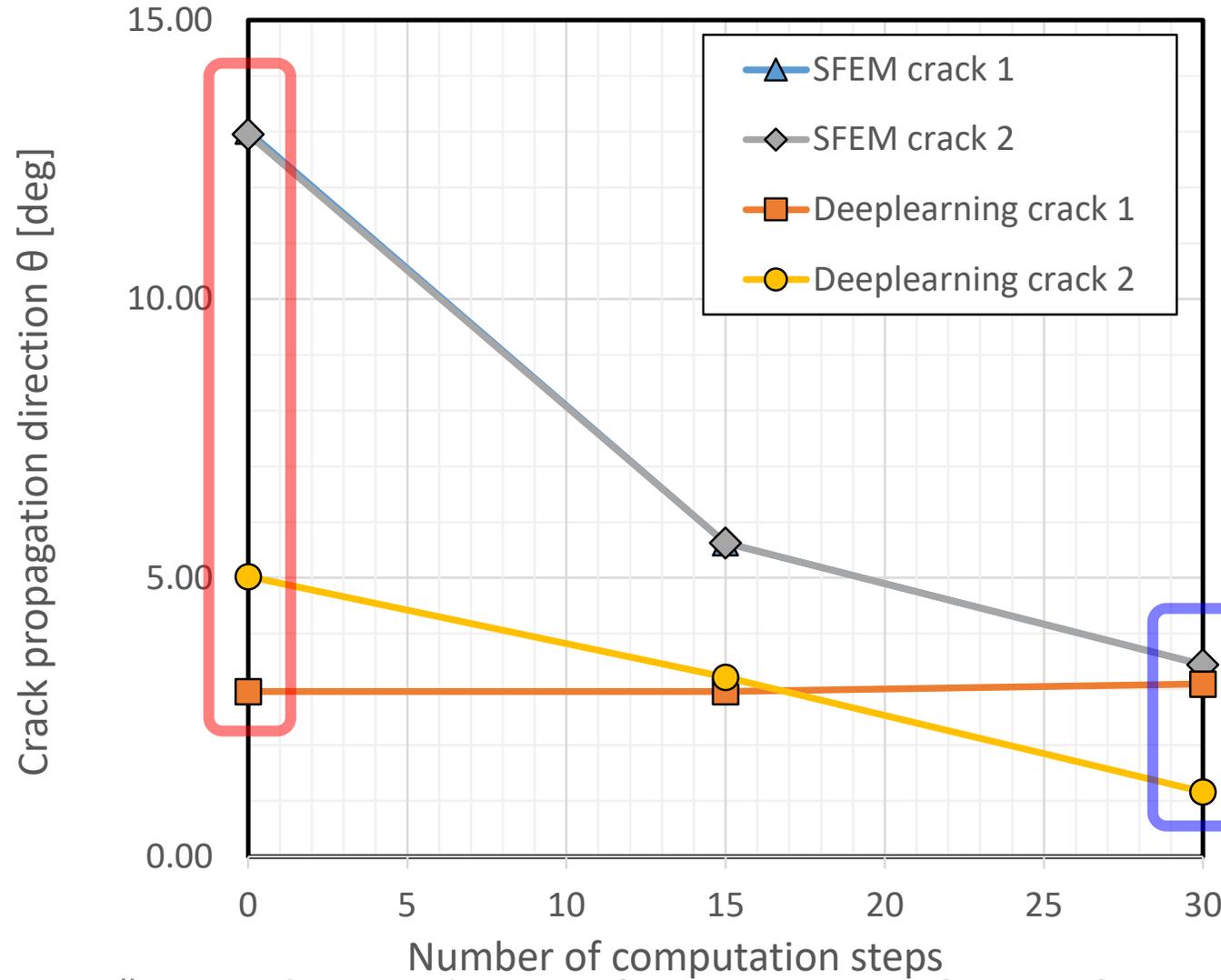


Crack propagation direction

15th step

initial

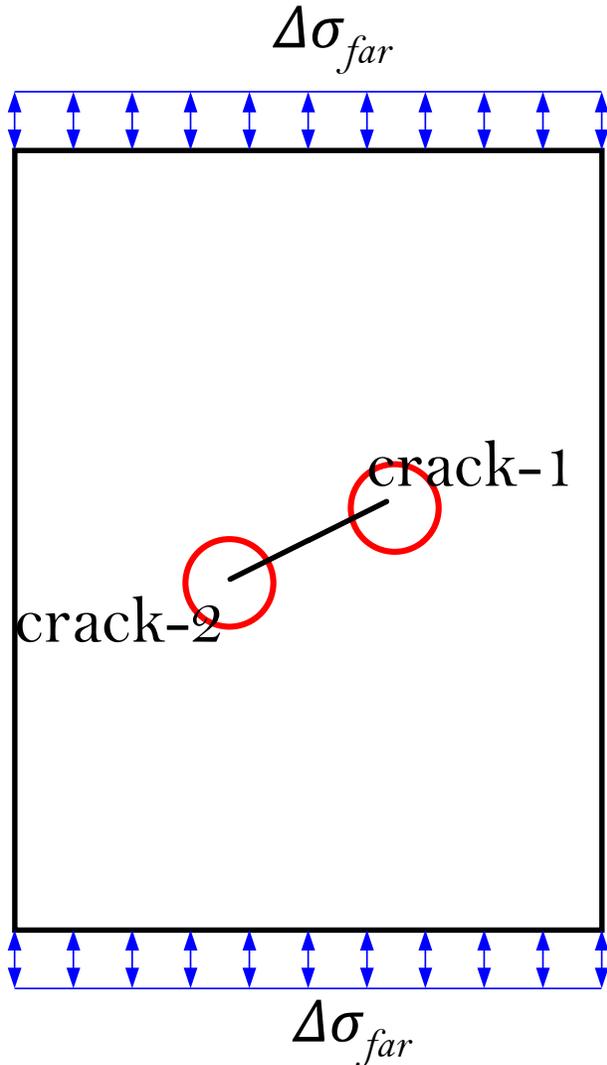
30th step



Not enough experience to get the appropriate prediction



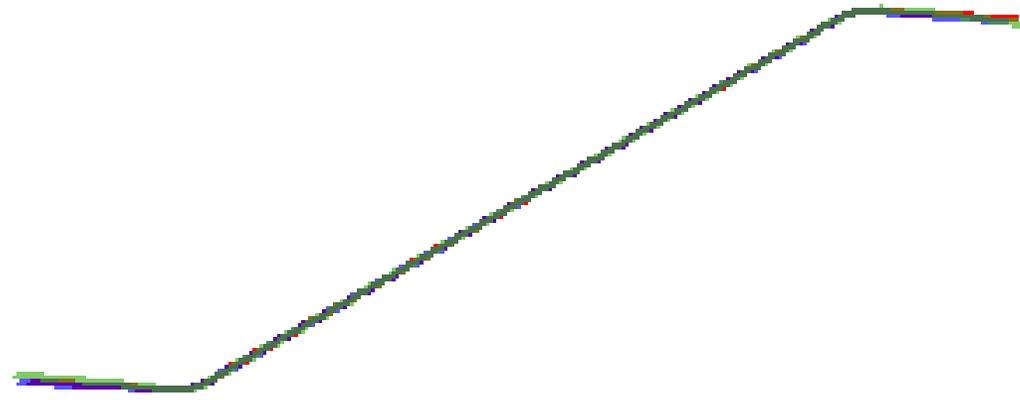
Numerical comparison at first computation step



	Crack position n	Crack prop. Vector: dP		Number of cycles: N
		dPx	dPy	
S-FEM	crack 1	0.974	-0.225	3.21×10^9
	crack 2	-0.974	0.224	3.21×10^9
Deeplearning	crack 1	0.967	-0.050	$2.65 \times 10^{8.68}$
	crack 2	-0.955	0.084	$3.84 \times 10^{8.98}$



Comparison between Phase 1, 2 and 3



Phase1: S.I.F.

Phase2: Stresses at crack tip

Phase3: Crack tip position and direction



Training conditions: latest trial

Training

Number of training sets

: 2860x5 (data augmented)

Initial crack angle (every 2.5°)

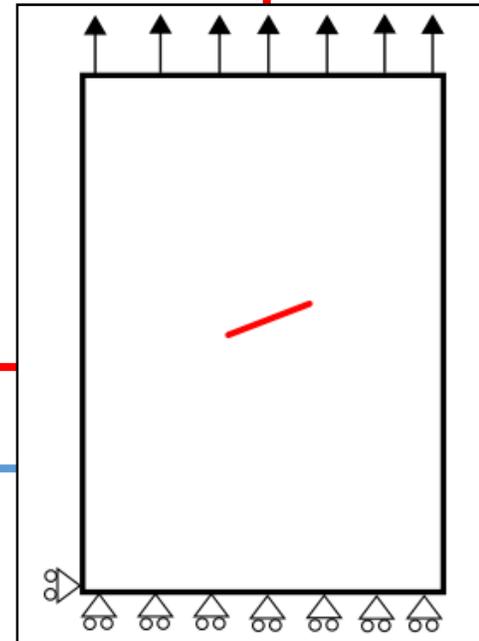
$\theta = 0^\circ$ to 87.5° without 45°

Hidden layers : 7 (125,75,30,15,30,75,125)

Optimizer : Adam

Verification model

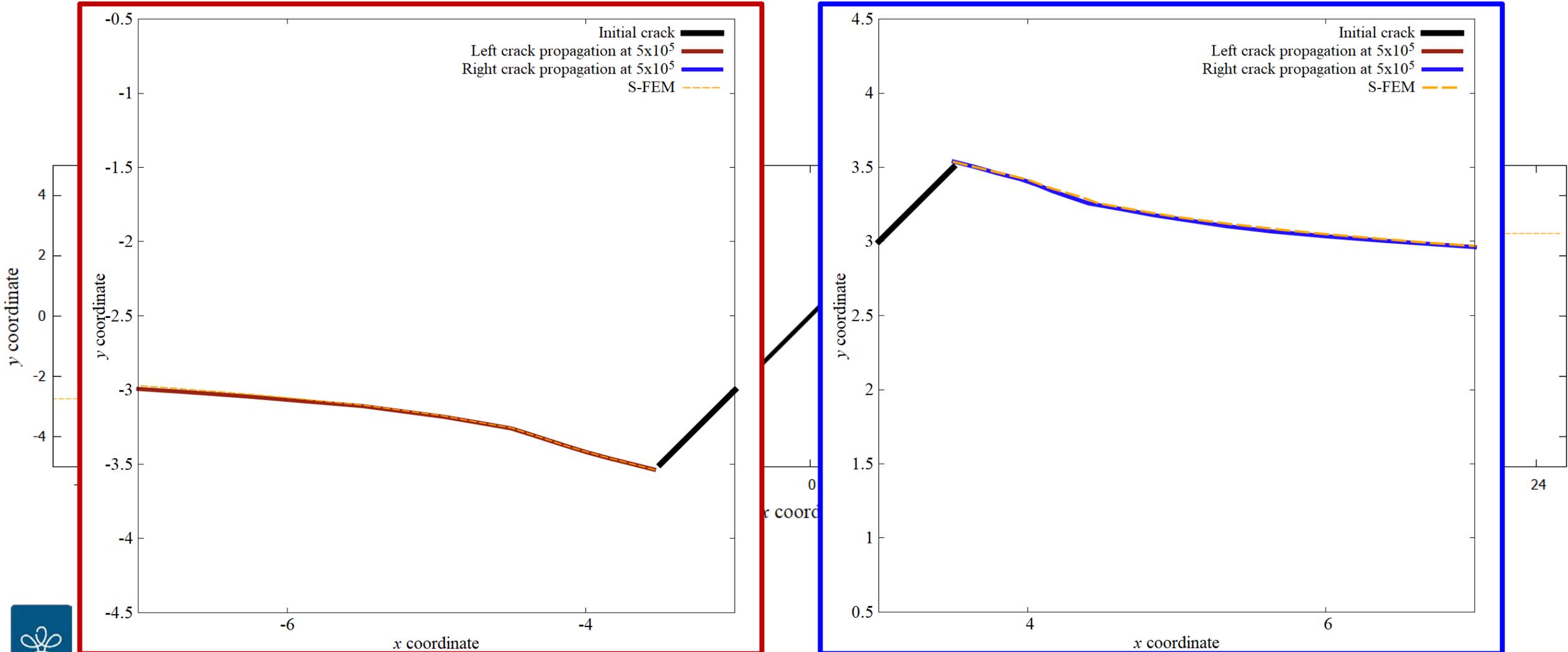
Initial crack angle 45°



Result of Phase 3

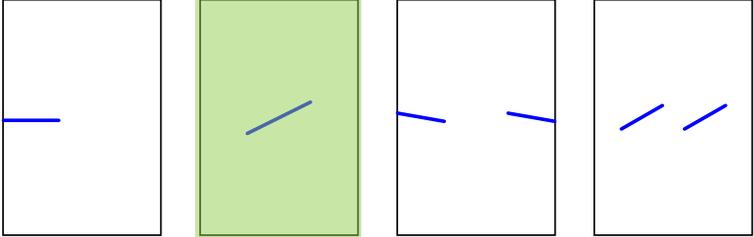
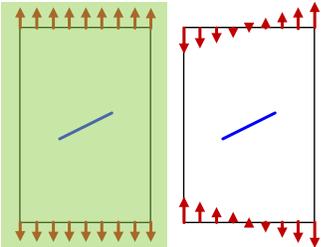
Num. of trainings: 5×10^5

Data augmentation to uniform density of data



Further work: extension for general crack prop.

Input

Interaction of two cracks	\mathbf{p}_{diff} \mathbf{d}_{diff} Relative crack posi. m Difference of direction vectors m	
Stress distribution	σ_b σ_{far} Bending stress MPa Stress MPa	
	da Crack prop. dist. m	\mathbf{p}_{crack} Crack tip position (Origin at crack center, vector)

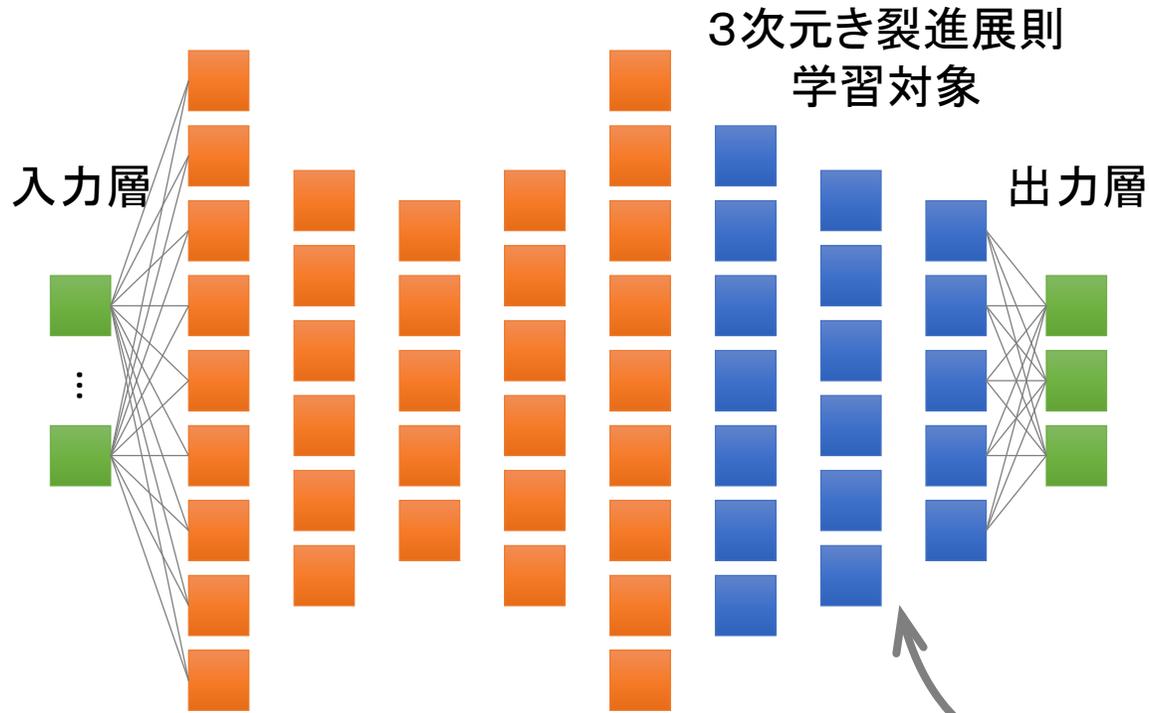
Output

All phases	da Crack prop. vector, dN cycles per a cal. step
------------	--



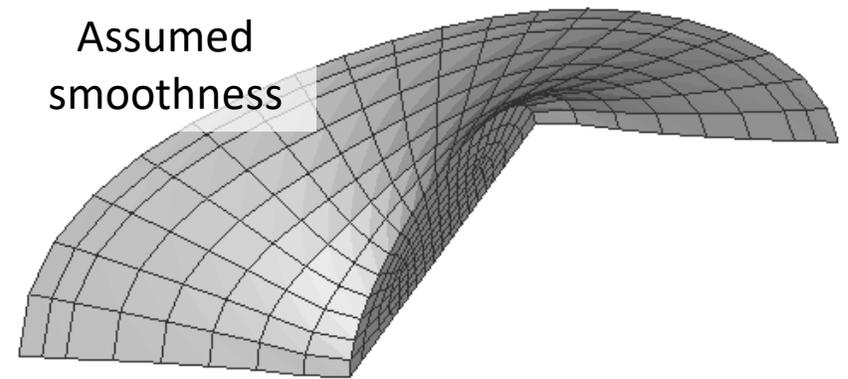
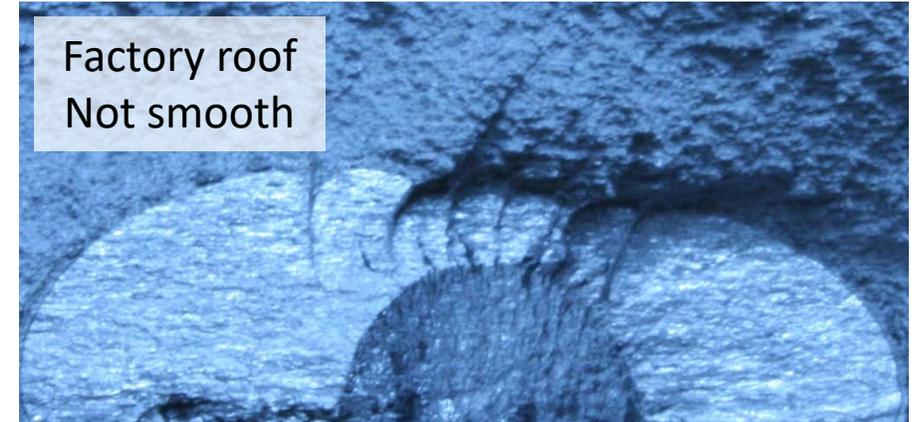
転移学習 transfer learning

2次元き裂進展から3次元へ



2次元き裂進展則
学習済ネットワーク
この部分は凍結

3次元き裂進展則
学習対象



K_{III} は形状に影響、実用上き裂深さと
き裂進展速度の低下の評価でよい



まとめ

- 原子力分野のみならず多くの分野で**超高精度を求め**すぎている
 - 問題の切り分けができていない
 - データの質についての議論ができていない
- より多くの**実例と有用性提示必要**
- 原子力発電プラントでは**質の高いデータが入手**できる環境、AI技術を使うのには**最も適した分野の1つ**
- 学習対象は**厳選必要**
 - 最適学習データの選別最も困難
 - **データ生成手法**の確立は最初に解決すべき



この講演で解説した機械学習に関するキーワード

- ニューラルネットワーク
 - MLP: Multi Layer Perceptron / CNN: Convolutional Neural Network
- ディープラーニング / 深層学習
- アクティベーション関数 / 活性化関数 (シグモイド、双曲線正接)
- バックプロパゲーション / 誤差逆伝播学習
- オプティマイゼーション / 最適化 (損失の低減方法)
- バッチノーマライゼーション、ドロップアウト
- オートエンコーダー
- データオーギュメンテーション / データ拡張 (水増し)
- トランスファートレーニング / 転移学習
- サロゲートモデル / 代替モデル
- GPU, Python, Anaconda, Theano, TensorFlow, Keras

