Privacy-Preserving Logistic Regression based on HEAAN Library

Andrey Kim[⋆], Miran Kim[†], Keewoo Lee[⋆], Yongsoo Song[⋆]

[⋆]Seoul National University

[†]University of California, San Diego

October 14, 2017

iDASH Privacy & Security Workshop 2017

Table of contents

Task3: HE based Logistic Regression Model Learning

- Logistic Regression
  - Regression model for categorical dependent variable.
  - Find a machine learning model for disease prediction using the genotype/phenotype data.
- Implementation Results

Iteration	**Learning**	**AUC**
7	10.25min	0.709
14	63.85min	0.715

#**(Features)** = 19, #**(Samples)** = 1421

HEAAN* Library

- Homomorphic Encryption for Arithmetic of Approximate Numbers [CKKS, to appear in AC'17]
- Efficient HE scheme over the real/complex numbers.

* 慧眼: insight, prescience

HEAAN* Library

- Homomorphic Encryption for Arithmetic of Approximate Numbers [CKKS, to appear in AC'17]
- Efficient HE scheme over the real/complex numbers.
- Provide an open-source implementation in C++ language. Available at github (https://github.com/kimandrik/HEAAN).

* 慧眼: insight, prescience

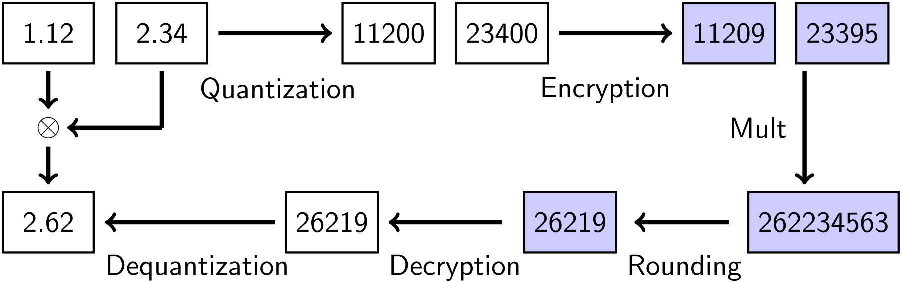
HEAAN* Library

- Homomorphic Encryption for Arithmetic of Approximate Numbers [CKKS, to appear in AC'17]
- Efficient HE scheme over the real/complex numbers.
- Provide an open-source implementation in C++ language. Available at github (https://github.com/kimandrik/HEAAN).
- Follow-up researches in applications
  - Encrypted Controller of Cyber-Physical System [KLSCKKS, IFAC'16]
  - Privacy-Preserving Logistic Regression [KSWXJ17, preprint]

* 慧眼: insight, prescience

Construction of HEAAN

- A noise of (Ring) LWE is considered to be a part of computation error.
- $\text{ct} = Enc(m) \quad \Rightarrow \quad Dec(\text{ct}) = m + e$ for some small e.
  - An approximate value can replace the original message.
- Support arithmetic operations and rounding of encrypted data.
- Linear bitsize of ctx modulus on the depth of a circuit & bit precision



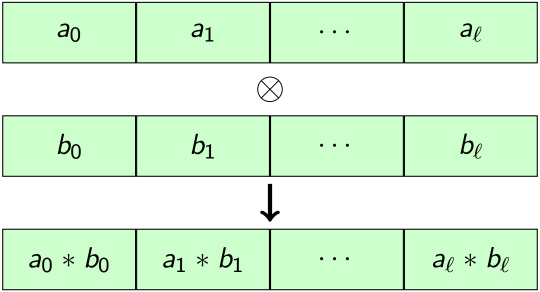
Functionality of HEAAN

- RLWE based construction over the cyclotomic ring $\mathcal{R} = \mathbb{Z}[X]/(X^N + 1)$ of a power-of-two dimension N.
- Packing multiple numbers (max. $N/2$) in a single ciphertext.

a_0	a_1	$\cdots$	a_ℓ

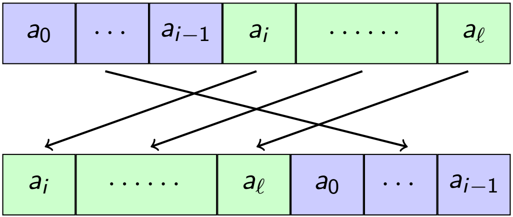
Functionality of HEAAN

- RLWE based construction over the cyclotomic ring $\mathcal{R} = \mathbb{Z}[X]/(X^N + 1)$ of a power-of-two dimension N.
- Packing multiple numbers (max. $N/2$) in a single ciphertext.
- Addition, multiplication, and rounding in a SIMD manner.



Functionality of HEAAN

- RLWE based construction over the cyclotomic ring $\mathcal{R} = \mathbb{Z}[X]/(X^N + 1)$ of a power-of-two dimension N.
- Packing multiple numbers (max. $N/2$) in a single ciphertext.
- Addition, multiplication, and rounding in a SIMD manner.
- Rotation on plaintext slots.



Database Encoding

User	Class (± 1)	Feature Vector ($\mathbb{R}^d$)			
1	$y_1 = 1$	$x_{11} = 2.78$	$x_{12} = 1.12$	$\cdots$	$x_{1d} = 3.05$
2	$y_2 = -1$	$x_{21} = 0.56$	$x_{22} = 1.58$	$\cdots$	$x_{2d} = 2.95$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	$y_n = 1$	$x_{n1} = 1.22$	$x_{n2} = 2.01$	$\cdots$	$x_{nd} = 4.31$

$$\Longrightarrow \quad Z = \begin{bmatrix} 1 & 2.78 & 1.12 & \cdots & 3.05 \\ -1 & -0.56 & -1.58 & \cdots & -2.95 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1.22 & 2.01 & \cdots & 4.31 \end{bmatrix} \in \mathbb{R}^{n \times (d+1)}$$

- Write $\vec{z}_i = y_i \cdot (1, \vec{x}_i) \in \mathbb{R}^{d+1}$ for $i = 1, 2, \ldots, n$.
- Generate an encryption of $Z = (\vec{z}_i)_i$.

Logistic Regression and Gradient Descent

- Find a modeling vector $\vec{\beta} \in \mathbb{R}^{d+1}$ which maximizes the likelihood

$$\prod_{i=1}^{n} \Pr[y_i|\vec{x}_i] \quad \text{for} \quad \Pr[y_i|\vec{x}_i] = \frac{1}{1+\exp(-\vec{z}_i \cdot \vec{\beta})}.$$

- Loss function $J(\vec{\beta}) = \frac{1}{n}\sum_{i=1}^{n} \log(1+\exp(-\vec{z}_i \cdot \vec{\beta}))$.
- $\bigtriangledown J(\vec{\beta}) = -\frac{1}{n}\sum_{i=1}^{n} \sigma(\vec{z}_i \cdot \vec{\beta}) \cdot \vec{z}_i$ for the sigmoid function σ.
- Repeat $\vec{\beta}_t \leftarrow \vec{\beta}_{t-1} - \alpha_t \cdot \bigtriangledown J(\vec{\beta}_{t-1})$ for $t = 1, 2, \cdots$.

Evaluation Strategy

$$\begin{array}{|c|c|c|c|} \hline \beta_0 & \beta_1 & \cdots & \beta_d \\ \hline \beta_0 & \beta_1 & \cdots & \beta_d \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \beta_0 & \beta_1 & \cdots & \beta_d \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|} \hline z_{10} & z_{11} & \cdots & z_{1d} \\ \hline z_{20} & z_{21} & \cdots & z_{2d} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline z_{n0} & z_{n1} & \cdots & z_{nd} \\ \hline \end{array}$$

Evaluation Strategy

$$
\begin{bmatrix}
\beta_0 & \beta_1 & \cdots & \beta_d \\
\beta_0 & \beta_1 & \cdots & \beta_d \\
\vdots & \vdots & \ddots & \vdots \\
\beta_0 & \beta_1 & \cdots & \beta_d
\end{bmatrix}
\otimes
\begin{bmatrix}
z_{10} & z_{11} & \cdots & z_{1d} \\
z_{20} & z_{21} & \cdots & z_{2d} \\
\vdots & \vdots & \ddots & \vdots \\
z_{n0} & z_{n1} & \cdots & z_{nd}
\end{bmatrix}
\rightarrow
\begin{bmatrix}
z_{10}\beta_0 & z_{11}\beta_1 & \cdots & z_{1d}\beta_d \\
z_{20}\beta_0 & z_{21}\beta_1 & \cdots & z_{2d}\beta_d \\
\vdots & \vdots & \ddots & \vdots \\
z_{n0}\beta_0 & z_{n1}\beta_1 & \cdots & z_{nd}\beta_d
\end{bmatrix}
$$

Evaluation Strategy

$$\begin{bmatrix} \beta_0 & \beta_1 & \cdots & \beta_d \\ \beta_0 & \beta_1 & \cdots & \beta_d \\ \vdots & \vdots & \ddots & \vdots \\ \beta_0 & \beta_1 & \cdots & \beta_d \end{bmatrix} \otimes \begin{bmatrix} z_{10} & z_{11} & \cdots & z_{1d} \\ z_{20} & z_{21} & \cdots & z_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n0} & z_{n1} & \cdots & z_{nd} \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} z_{10}\beta_0 & z_{11}\beta_1 & \cdots & z_{1d}\beta_d \\ z_{20}\beta_0 & z_{21}\beta_1 & \cdots & z_{2d}\beta_d \\ \vdots & \vdots & \ddots & \vdots \\ z_{n0}\beta_0 & z_{n1}\beta_1 & \cdots & z_{nd}\beta_d \end{bmatrix} \xrightarrow{\text{Rotate \& Sum}} \begin{bmatrix} \vec{z}_1 \cdot \vec{\beta} & \star & \cdots & \star \\ \vec{z}_2 \cdot \vec{\beta} & \star & \cdots & \star \\ \vdots & \vdots & \ddots & \vdots \\ \vec{z}_n \cdot \vec{\beta} & \star & \cdots & \star \end{bmatrix}$$

Evaluation Strategy

z_{10}	z_{11}	$\cdots$	z_{1d}
z_{20}	z_{21}	$\cdots$	z_{2d}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
z_{n0}	z_{n1}	$\cdots$	z_{nd}

$\vec{z}_1 \cdot \vec{\beta}$	$\star$	$\cdots$	$\star$
$\vec{z}_2 \cdot \vec{\beta}$	$\star$	$\cdots$	$\star$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\vec{z}_n \cdot \vec{\beta}$	$\star$	$\cdots$	$\star$

Evaluation Strategy

$$\begin{array}{|c|c|c|c|}\hline z_{10} & z_{11} & \cdots & z_{1d} \\ \hline z_{20} & z_{21} & \cdots & z_{2d} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline z_{n0} & z_{n1} & \cdots & z_{nd} \\ \hline \end{array}$$

$$\begin{array}{|c|c|c|c|}\hline \vec{z}_1 \cdot \vec{\beta} & \vec{z}_1 \cdot \vec{\beta} & \cdots & \vec{z}_1 \cdot \vec{\beta} \\ \hline \vec{z}_2 \cdot \vec{\beta} & \vec{z}_2 \cdot \vec{\beta} & \cdots & \vec{z}_2 \cdot \vec{\beta} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \vec{z}_n \cdot \vec{\beta} & \vec{z}_n \cdot \vec{\beta} & \cdots & \vec{z}_n \cdot \vec{\beta} \\ \hline \end{array} \xleftarrow{\text{Replicate}} \begin{array}{|c|c|c|c|}\hline \vec{z}_1 \cdot \vec{\beta} & \star & \cdots & \star \\ \hline \vec{z}_2 \cdot \vec{\beta} & \star & \cdots & \star \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \vec{z}_n \cdot \vec{\beta} & \star & \cdots & \star \\ \hline \end{array}$$

Evaluation Strategy

$$\begin{array}{|c|c|c|}\hline \sigma(-\vec{z}_1 \cdot \vec{\beta}) & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \\ \hline \sigma(-\vec{z}_2 \cdot \vec{\beta}) & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \\ \hline \vdots & \ddots & \vdots \\ \hline \sigma(-\vec{z}_n \cdot \vec{\beta}) & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \\ \hline \end{array}$$

$$\begin{array}{|c|c|c|c|}\hline z_{10} & z_{11} & \cdots & z_{1d} \\ \hline z_{20} & z_{21} & \cdots & z_{2d} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline z_{n0} & z_{n1} & \cdots & z_{nd} \\ \hline \end{array}$$

Sigmoid $\uparrow$

$$\begin{array}{|c|c|c|c|}\hline \vec{z}_1 \cdot \vec{\beta} & \vec{z}_1 \cdot \vec{\beta} & \cdots & \vec{z}_1 \cdot \vec{\beta} \\ \hline \vec{z}_2 \cdot \vec{\beta} & \vec{z}_2 \cdot \vec{\beta} & \cdots & \vec{z}_2 \cdot \vec{\beta} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \vec{z}_n \cdot \vec{\beta} & \vec{z}_n \cdot \vec{\beta} & \cdots & \vec{z}_n \cdot \vec{\beta} \\ \hline \end{array} \xleftarrow{\text{Replicate}} \begin{array}{|c|c|c|c|}\hline \vec{z}_1 \cdot \vec{\beta} & \star & \cdots & \star \\ \hline \vec{z}_2 \cdot \vec{\beta} & \star & \cdots & \star \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \vec{z}_n \cdot \vec{\beta} & \star & \cdots & \star \\ \hline \end{array}$$

Evaluation Strategy

$\sigma(-\vec{z}_1 \cdot \vec{\beta})$	$\cdots$	$\sigma(-\vec{z}_1 \cdot \vec{\beta})$
$\sigma(-\vec{z}_2 \cdot \vec{\beta})$	$\cdots$	$\sigma(-\vec{z}_1 \cdot \vec{\beta})$
$\vdots$	$\ddots$	$\vdots$
$\sigma(-\vec{z}_n \cdot \vec{\beta})$	$\cdots$	$\sigma(-\vec{z}_1 \cdot \vec{\beta})$

z_{10}	z_{11}	$\cdots$	z_{1d}
z_{20}	z_{21}	$\cdots$	z_{2d}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
z_{n0}	z_{n1}	$\cdots$	z_{nd}

Evaluation Strategy

$$\begin{bmatrix} \sigma(-\vec{z}_1 \cdot \vec{\beta}) & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \\ \sigma(-\vec{z}_2 \cdot \vec{\beta}) & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \\ \vdots & \ddots & \vdots \\ \sigma(-\vec{z}_n \cdot \vec{\beta}) & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \end{bmatrix} \otimes \begin{bmatrix} z_{10} & z_{11} & \cdots & z_{1d} \\ z_{20} & z_{21} & \cdots & z_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n0} & z_{n1} & \cdots & z_{nd} \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} \sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{10} & \cdots & \sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{1d} \\ \sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{20} & \cdots & \sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{2d} \\ \vdots & \ddots & \vdots \\ \sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{n0} & \cdots & \sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{nd} \end{bmatrix}$$

Evaluation Strategy

$\sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{10}$	$\cdots$	$\sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{1d}$
$\sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{20}$	$\cdots$	$\sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{2d}$
$\vdots$	$\ddots$	$\vdots$
$\sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{n0}$	$\cdots$	$\sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{nd}$

Evaluation Strategy

$\sum_{i=1}^{n} \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{i0}$	$\cdots$	$\sum_{i=1}^{n} \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{id}$
$\sum_{i=1}^{n} \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{i0}$	$\cdots$	$\sum_{i=1}^{n} \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{id}$
$\vdots$	$\ddots$	$\vdots$
$\sum_{i=1}^{n} \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{i0}$	$\cdots$	$\sum_{i=1}^{n} \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{id}$

Rotate & Sum

$\sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{10}$	$\cdots$	$\sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{1d}$
$\sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{20}$	$\cdots$	$\sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{2d}$
$\vdots$	$\ddots$	$\vdots$
$\sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{n0}$	$\cdots$	$\sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{nd}$

Evaluation Strategy

$\sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{10}$	$\cdots$	$\sigma(-\vec{z}_1 \cdot \vec{\beta}) \cdot z_{1d}$
$\sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{20}$	$\cdots$	$\sigma(-\vec{z}_2 \cdot \vec{\beta}) \cdot z_{2d}$
$\vdots$	$\ddots$	$\vdots$
$\sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{n0}$	$\cdots$	$\sigma(-\vec{z}_n \cdot \vec{\beta}) \cdot z_{nd}$

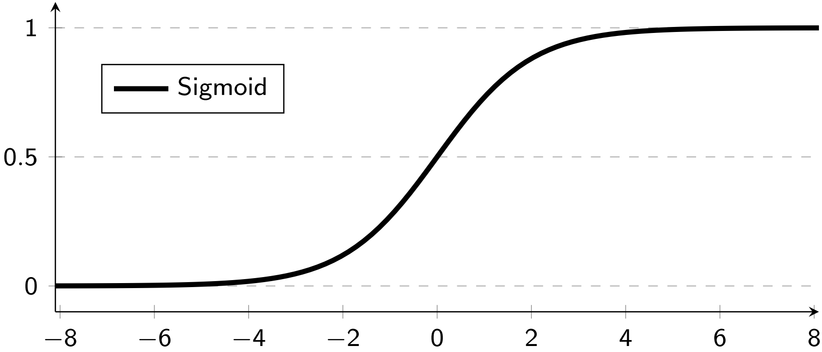
Rotate & Sum

$\sum_{i=1}^n \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{i0}$	$\cdots$	$\sum_{i=1}^n \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{id}$
$\sum_{i=1}^n \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{i0}$	$\cdots$	$\sum_{i=1}^n \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{id}$
$\vdots$	$\ddots$	$\vdots$
$\sum_{i=1}^n \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{i0}$	$\cdots$	$\sum_{i=1}^n \sigma(-\vec{z}_i \cdot \vec{\beta}) \cdot z_{id}$

Update

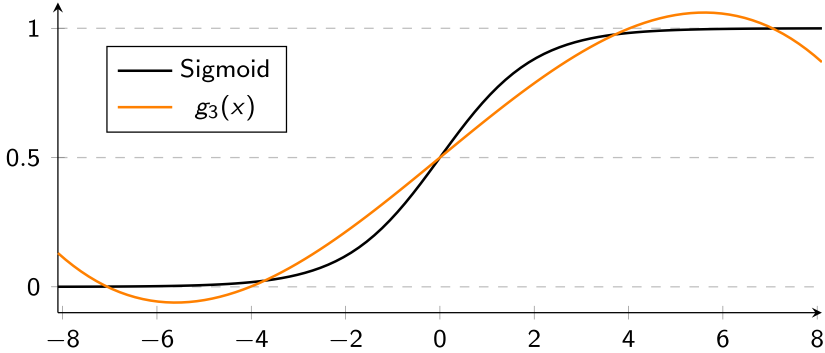
β_0	β_1	$\cdots$	β_d
β_0	β_1	$\cdots$	β_d
$\vdots$	$\vdots$	$\ddots$	$\vdots$
β_0	β_1	$\cdots$	β_d

Sigmoid Function



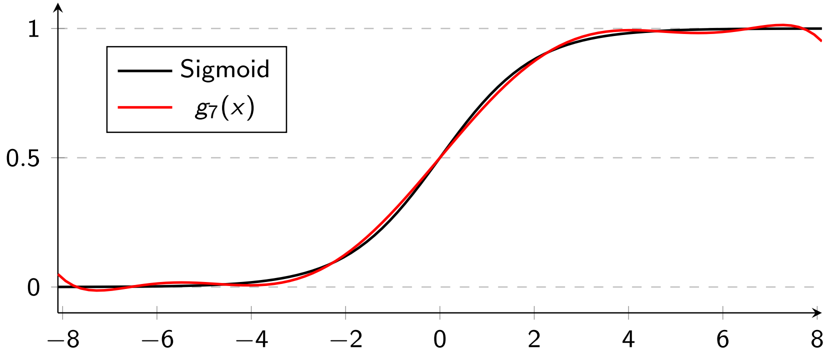
- The most widely used activation function.
- Use the least squares on [-8,8] to find an approximate polynomial.

Degree 3



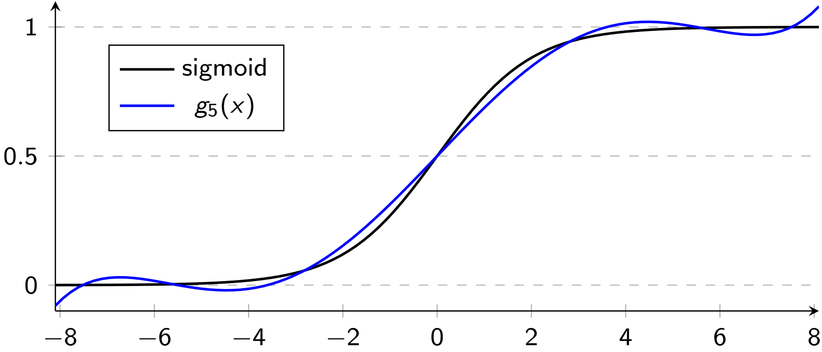
- Pros: Small depth
- Cons: Maximum error > 0.1.

Degree 7



- Pros: Maximum error ≈ 0.03
- Cons: Computation cost.

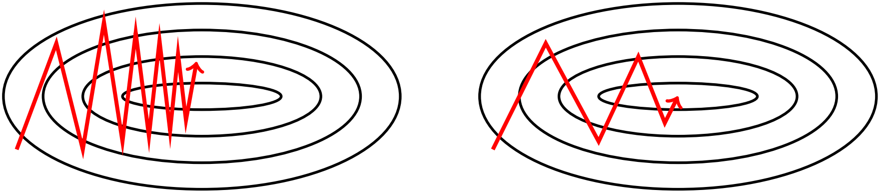
Degree 5



- Maximum error ≈ 0.06.

Momentum Method

- Momentum is a method that accelerates GD in the relevant direction and dampens oscillations.
- Add a fraction of the update vector of the past time step to the current update vector.



Standard GD Momentum method

Nesterov Accelerated Gradient

- Guarantee a better rate of convergence $O(1/t^2)$ (vs. $O(1/t)$)
- Evaluate the gradient at this "looked-ahead" position.

$$\vec{v}_t = \gamma \cdot \vec{v}_{t-1} + \alpha_t \cdot \bigtriangledown J(\vec{\beta}_{t-1} - \gamma \cdot \vec{v}_{t-1})$$
$$\vec{\beta}_t = \vec{\beta}_{t-1} - \vec{v}_t$$

Experimental Results

Intel(R) Xeon(R) CPU E5-2670 v2 @ 2.50GHz processor.

Parameters				Results					
Iter	**Deg**	$\log N$	$\log q$	**Enc Time**	**Learn Time**	**PK Size**	**Encrypted Database**	**Correctness**	**AUC**
7	5	16	1229	0.95min	10.25min	1.2GB	38 MB	815 / 1421	0.709
14	5	17	2405	2.33min	63.85min	2.4GB	75 MB	888 / 1421	0.715

#**(Features)** = 19, #**(Samples)** = 1421


감사합니다