Тооцон бодох Математик: Тооцооллын Инженер

Содержание

Өмнөх Үг

Өнгөрсөн, одоо, ба ирээдүйгээр аялуулагч нь утга, уран зохиол болон урлаг юм болов уу. Өнгөрсөн, одоо, ба ирээдүйг магадлан тайлбарлагч нь шинжлэх ухаан ба инженерчилэл юм болов уу.

Отгонбаатар Соронзоболд 24 мая 2020 г.

PART

I

Хүн-Машин

Remark (Л. Түдэв, Оройгүй Сүм). Урьдын баян хожим үгүйрвэл хөөрхий
Урьдын ноён хожим буурвал хөөрхий
Урин тэнэгүүдийн дунд ганц цэцэн хөөрхий

1 Лагранж интерполяц болон хамгийн бага квадратын арга

Интерполяц гэдэг нь тооцон бодох математикийн (Computational Mathematics) хувьд өгөгдсөн тасралттай өгөгдөлд (data) зориулан модел (model) буюу хамаарлын тэгшитгэлийг боловсруулах арга бөгөөд тухайн тэгшитгэлийг ашиглан өгөгдсөн өгөгдөлийн муж дотроос дурын өгөгдөл бий болгох арга юм. Нэгэн жишээ авч үзээ.

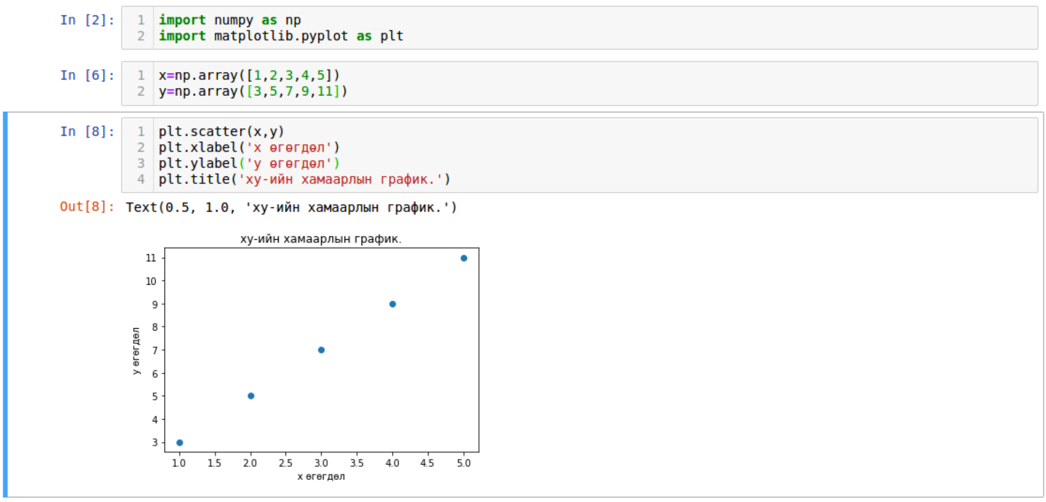


Рис. 1.

Рис. 2. "Лодонгийн Түдэв - Монгол Улсын хөдөлмөрийн баатар, Ардын Уран Зохиолч, СГЗ, Төрийн шагналт, доктор, УБИС-ийн хүндэт профессор, XX зууны манлай сэтгүүлч. 1935 оны 2 сарын 9-нд Говь-Алтай аймгийн Наран (одоогийн Халиун) сумын нутаг "Өвөр замын бууц" гэдэг газар, малчин айлын гэр бүлд төржээ."

Энэ жишээний хувьд x-ийн утгын муж нь 1-ээс 5-ын хооронд орших бөгөөд үүнд харгалзах y-ын хамаарлын утгыг үзүүлсэн байна. Тэгвэл бидний дараагийн асуулт маань x-ийн утга болох $(1, 5)$-ын хооронд орших дурын өгөгдөлийн хувьд харгалзах y-ын утгыг хэрхэн олох вэ? (тухайлбал дурын өгөгдөлд $x = 2.6$ харгалзах y юу байх вэ?). Зурагт үзүүлсэн xy-ын хамаарлыг анхааралтай ажиглабал

$$f(x_i) = y_i = 1 + 2 \cdot x_i \tag{1.1}$$

гэдгийг хялбархан харж болно. Өөрөөр хэлбэл энэхүү хамаарлын тэгшитгэлийг бидэнд өгөгдсөн өгөгдөлд зориулсан модел (Шугаман модел) гэж нэрлэдэг. Өгөгдсөн өгөгдөлийн муж дотроос (зөвхөн муж дотроос гэдгийг анхаараарай) дурын өгөгдөлийг үүсгэх хамаарлын тэгшитгэлийг олж буй энэхүү аргыг дээр дурдсанчлан интерполяцлах гэдэг. Бидний авч үзэж буй жишээний хувьд өгөгдөл маань шугаман хамааралтай учир (x, y) хамаарлын тэгшитгэлийг олоход амархан байлаа. Харин шугаман бус хамааралтай өгөгдөлүүдийн хувьд хамаарлын тэгшитгэлийг олоход хэцүү байдаг төдийгүй хамаарлын тэгшитгэлийг нь байгуулах олон арга байдаг. Иймээс энэхүү хэсэгт бид зөвхөн Лагранжийн интерполяц хэмээх аргыг авч үзэх болно.

2 Лагранж интерполяц

Өгөгдсөн өгөгдөлийн хувьд $(x_i, y_i), \quad i = 1, 2, \ldots, N$, хамаарлын функцыг $f(x)$ олох нь гол зорилго юм.Өөрөөр хэлбэл $f(x_i) = y_i$ буюу дээр дурдсан жишээг харна уу. Тэгвэл хайж буй функцыг доорх хэлбэртэй бичие,

$$f(x_i) = \sum_{k=1}^{M} \phi_k(x_i)\alpha_k, \tag{2.1}$$

энд суурь функц (base function) $\phi_k(x_i)$ ба параметер α_k. Дээр дурдсан тэгшитгэл (1.1)-ийн хувьд $M = 2, \quad i = 1, 2, \ldots, 5$ ба $\phi_1(x_i) = 1, \quad \phi_2(x_i) = x_i$ бөгөөд түүнчлэн тэгшитгэл (2.1),

$$f(x_i) = y_i = \phi_1(x_i)\alpha_1 + \phi_2(x_i)\alpha_2 = \alpha_1 + x_i\alpha_2, \tag{2.2}$$

хэлбэртэй болох ба хэрэв дээрх тэгшигтгэлээс α_k параметрийг олбол $\alpha_1 = 1$ мөн $\alpha_2 = 2$ болох буюу тэгшитгэл (1.1) ийг амархан тодорхойлж болно. Цаашиалбал суурь функц маань доорх 2 нөхцлийг

$$\phi_k(x_k) = 1, \quad \phi_k(x_i) = 0, (k = 1, 2, \ldots N) \tag{2.3}$$

хангадаг гэж үзэе. Тэгвэл

$$\phi_k(x) = \frac{(x - x_1)(x - x_2)\ldots(x - x_N)}{(x_k - x_1)(x_k - x_2)\ldots(x_k - x_N)} \tag{2.4}$$

хэлбэртэйгээр суурь функцыг бичих боломжтой. Цаашид ойлгомжтой байхын тулд $\phi_k(x)$-ийг $l_k(x)$ гэж тэмдэглэнэ.

$$f(x) = \sum_{k=1}^{N} l_k(x)\alpha_k, \tag{2.5}$$

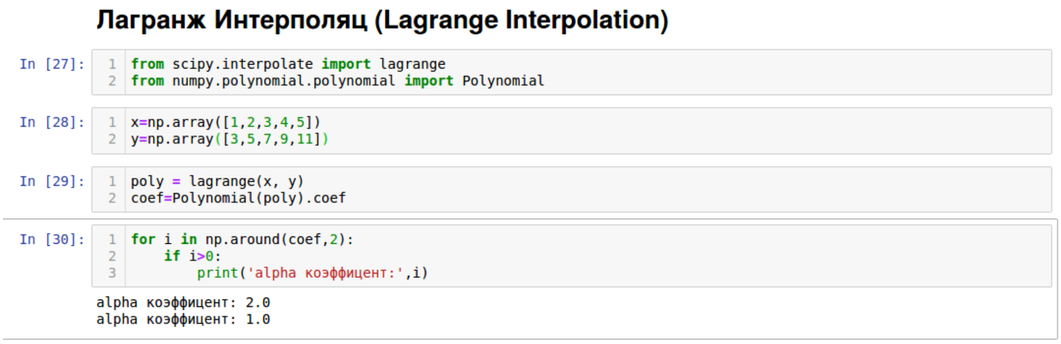
буюу параметр α_k-г олбол,

$$f(x_i) = y_i = \sum_{k=1}^{M} l_k(x_i)\alpha_k = \alpha_i, \tag{2.6}$$

цаашилбал

$$f(x) = \sum_{k=1}^{N} l_k(x)y_k. \tag{2.7}$$

Одоо дээр дурдсан жишээний хувьд Лагранж интерполяц (2.7) ашиглан коэффициентыг нь тодорхойлоё. Цаашилбал бид энэхүү аргад зориулсан бэлэн python програмчилалын хэл ашиглан тооцоогоо хялбархан гүйцэтгэе. Эндээс



Лагранж Интерполяц (Lagrange Interpolation)

In [27]:
```
from scipy.interpolate import lagrange
from numpy.polynomial.polynomial import Polynomial
```

In [28]:
```
x=np.array([1,2,3,4,5])
y=np.array([3,5,7,9,11])
```

In [29]:
```
poly = lagrange(x, y)
coef=Polynomial(poly).coef
```

In [30]:
```
for i in np.around(coef,2):
    if i>0:
        print('alpha коэффицент:',i)
```

```
alpha коэффицент: 2.0
alpha коэффицент: 1.0
```

Рис. 3.

харвал бидний бидний таамаглаж олсон шугаман тэгшитгэл (1.1) болон (2.2) дээр олсон үр дүнгүүдтэй ижил байна. Өмнө үзүүлснээр бэлэн Python package ашиглахаас гадна, бид тэгшитгэл (2.7)-д зориулан өөрсдийн python кодыг маш хялбархан бичиж болох буюу доорх байдлаар бичигдэнэ. Өөр нэгэн жишээнд

```
In [66]:
def lk(x,y,z):
    s=0
    for k in range(0,len(x)):
        m=1
        for i in range(0,len(x)):
            if i!=k:
                m=m*(z-x[i])/(x[k]-x[i])
            else:
                continue
        s=s+y[k]*m
    return s
```

Рис. 4.

өөрсдийн бичсэн дээрх кодыг ашиглаж үзээ. Тэгвэл (x, y) өгөгдөл өгөгдсөн үед дурын завсрын MiddlePoints өгөгдөлд харгалзах y-ын утгыг доорх байдлаар тодорхойллоо.

```
In [79]: print(x, x.shape)
[-1  0  1  2  3  4  5  6  7  8  9] (11,)

In [80]: print(y, y.shape)
[ -1   0   1   8  27  64 125 216 343 512 729] (11,)

In [86]: MiddlePoins=np.array(MiddlePoins)
         y_estimated=np.array(LagrangeFunction)

In [87]: print(MiddlePoins)
[-0.5  0.5  1.5  2.5  3.5  4.5  5.5  6.5  7.5  8.5  9.5]

In [88]: print(y_estimated)
[-1.25000e-01  1.25000e-01  3.37500e+00  1.56250e+01  4.28750e+01
  9.11250e+01  1.66375e+02  2.74625e+02  4.21875e+02  6.14125e+02
  8.57375e+02]

In [89]: plt.scatter(MiddlePoins,y_estimated)
         plt.xlabel('MiddlePoins өгөгдөл')
         plt.ylabel('LagrangeFunction өгөгдөл')
         plt.title('(MiddlePoins,LagrangeFunction)-ийн хамаарлын график.')
Out[89]: Text(0.5, 1.0, '(MiddlePoins,LagrangeFunction)-ийн хамаарлын график.')
```

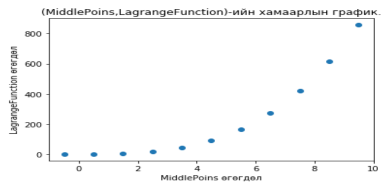


Рис. 5.

PART

II

Хүн-Машин

> Remark (Л. Түдэв, Оройгүй Сүм). Нумгүй анчин ав дээр гэмшил
> Номгүй лам хурал дээр гэмшил
> Үйлгүй хүүхэн хурим дээр гэмшил

3 Хамгийн бага квадратын арга (Least Squares Method)

Интерполяц гүйцэтгэх өгөгдөлийн тоо нь тодорхойлохыг зорьж буй параметрийн тооноос их байх үед хэрхэн тухайн параметрүүдийг олж болохыг сонирхоё. Бидэнд доорх жишээний параметрүүдийг олох шаардлагатай болсон гэж үзэе,

$$2x = b_1, \quad 3x = b_2, \quad 4x = b_3. \tag{3.1}$$

Энэхүү тэгштгэлийг матриц хэлбэрт бичие,

$$\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} x = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \tag{3.2}$$

Тэгшитгэлийн тоо $N = 3$ ба параметр нь $M = 1$ байна. Хэрэв $M = 3$ буюу тэгшитгэлийн тоо $N = 3$-нууд нь ижил байсан бол тэгшитгэл (3.2)-ийн шийд болох параметрүүдийг маш хялбархан олох боломжтой. Параметр ба тэгшитгэлийн тоо нь ижил биш үед шийд буюу параметрийг хайхын тулд хамгийн бага кватратын аргыг ашиглая,

$$e = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2. \tag{3.3}$$

Параметр x-ийн утгыг олохын тулд нь дээрх илэрхийллэлээс нэгдүгээр эрэмбийн уламжлал авсны дараа тэгшитгэл $\dfrac{de}{dx} = 0$- ийг тэгтэй тэнцүүлнэ. Тэгвэл,

$$2\left[(2x - b_1) \cdot 2 + (3x - b_2) \cdot 3 + (4x - b_3) \cdot 4\right] = 0, \tag{3.4}$$

буюу ойролцоо шийд нь

$$\bar{x} = \frac{2b_1 + 3b_2 + 4b_3}{2^2 + 3^2 + 4^2}. \tag{3.5}$$

Үүнээс үндэслэн ерөнхий шийдийг доорх хэлбэртэй бичиж болно

$$\vec{a}\bar{x} = \vec{b}, \quad \bar{x} = \frac{(\vec{a}^T\vec{b})}{(\vec{a}^T\vec{a})}. \tag{3.6}$$

Рис. 6. “Сэнгийн Эрдэнэ Хэнтий аймгийн Биндэр суманд 1929 оны 12 сарын 7-нд төрсөн. 1949 онд офицерын сургууль, 1955 онд МУИС төгсчээ. "Амьдралын тойрог "Баян бүрд "Занабазар "Малын хөлийн тоос "Наран тогоруу "Хойт насандаа учирна"зэрэг бүтээл хэвлүүлжээ. 1965 онд Монгол Улсын төрийн шагнал, 1976 онд МЗЭ-ийн шагнал, 1994 онд Ардын Уран Зохиолч цол хүртжээ.”

Дээрх жишээгээ өргөтгөж матриц хэлбэрээр илэрхийлбэл,

$$A\vec{x} = \vec{b}. \tag{3.7}$$

Хайж буй ойролцоо шийдийн $\bar{x}$ хувьд,

$$A\bar{x} = \vec{b}, \tag{3.8}$$

бөгөөд тэгшитгэл (3.6)-н дагуу

$$A^T A\bar{x} = A^T\vec{b}, \quad \bar{x} = (A^T A)^{-1}A^T\vec{b}. \tag{3.9}$$

3.1 Хамгийн бага кватратын арга: 2 хэмжээст өгөгдөлийн параметрийг олох

Доор үзүүлсэн 2 хэмжээст өгөгдөлийн (y_i, z_i) хоорондын хамаарлын параметрийг (C, D) олоё.

$$C + D \cdot y_i \approx z_i. \tag{3.10}$$

Дээрх тэгшитгэлийг матриц хэлбэрт бичвэл,

$$\begin{pmatrix} 1 & y_1 \\ 1 & y_2 \\ 1 & y_3 \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix}. \tag{3.11}$$

Энэхүү тэгшигэлийн өгөгдөлийн тоо нь параметрийн тоотой ижил биш учир шийдийг олохын тулд тэгшитгэл (3.9)-ийг хэрэглэн өгөгдөлийг 2x2 матриц хэлбэрт оруулна,

$$\begin{pmatrix} N & \sum_i^N y_i \\ \sum_i^N y_i & \sum_i^N y_i^2 \end{pmatrix} \begin{pmatrix} \bar{C} \\ \bar{D} \end{pmatrix} = \begin{pmatrix} \sum_i^N z_i \\ \sum_i^N y_i z_i \end{pmatrix}. \tag{3.12}$$

Ойролцоо шийд нь $(\bar{C}, \bar{D})$,

$$\begin{aligned} \bar{C} &= \frac{\sum_i^N y_i^2 \cdot \sum_i^N z_i - \sum_i^N y_i \sum_i^N y_i z_i}{N \sum_i^N y_i^2 - \left(\sum_i^N y_i\right)^2}, \\ \bar{D} &= \frac{N \sum_i^N y_i z_i - \sum_i^N y_i \sum_i^N z_i}{N \sum_i^N y_i^2 - \left(\sum_i^N y_i\right)^2}. \end{aligned} \tag{3.13}$$

3.1.1 Жишээ бодлого

In [1]:
```python
import numpy as np
import matplotlib.pyplot as plt
```

Хамгийн бага квадратын арга (Least Squares Method)

Өгөгдсөн өгөгдөл y-д харгалзах z-ын утга нь шугаман хамааралтай гэж үзэж доорх жишээг сонирхоё

$$y \in (0, 10), \quad z = 0.5 \cdot y - 1.$$

In [30]:
```python
y=[]
z=[]

for i in range(0,10):
    y.append(i)
    z.append(0.5*i-1)

y=np.array(y).reshape(10,1)
z=np.array(z).reshape(10,1)
```

Рис. 7.

Бидний зорилго бол дээр боловсруулсан аргыг ашиглан өгөгдсөн жишээний $(0.5, -1)$ утгыг олох юм. y-д харгалзах z өгөгдөл нь,

In [33]:
```python
print(z)
```

```
[[-1. ]
 [-0.5]
 [ 0. ]
 [ 0.5]
 [ 1. ]
 [ 1.5]
 [ 2. ]
 [ 2.5]
 [ 3. ]
 [ 3.5]]
```

Тэгшитгэл 12 ийн матрицийг A гэж тэмдэглэе

$$A = \begin{pmatrix} N & \sum_i^N y_i \\ \sum_i^N y_i & \sum_i^N y_i^2 \end{pmatrix}$$

бөгөөд харгалзах элементийн утгыг нь олоё.

In [38]:
```python
A=np.zeros([2,2])
A[0,0]=len(y)
A[0,1]=sum(y)
A[1,0]=sum(y)
A[1,1]=sum(y**2)
```

In [39]:
```python
print(A)
```

```
[[ 10.  45.]
 [ 45. 285.]]
```

Рис. 8.

Тэгшитгэл 12-ийн баруун талын матрицийг Y гэж тэмдэглэн харгалзах элементүүдийг нь олоё,

$$Y = \begin{pmatrix} \sum_i^N z_i \\ \sum_i^N y_i z_i \end{pmatrix}.$$

In [40]:
```python
s=0
for i in range(0,10):
    s+=y[i]*z[i]
```

In [41]:
```python
print(s)
```

```
[97.5]
```

In [42]:
```python
Y=np.zeros([2,1])
Y[0,0]=sum(z)
Y[1,0]=s
```

In [43]:
```python
print(Y)
```

```
[[12.5]
 [97.5]]
```

Рис. 9.

Одоо тэгшитгэл 9-ийн дагуу параметрийг $(\bar{C}, \bar{D})$ олоё.

$$res = \begin{pmatrix} \bar{C} \\ \bar{D} \end{pmatrix}.$$

In [25]:
```
from numpy.linalg import inv
inv_A=inv(A)
res=np.matmul(inv_A,Y)
```

In [26]:
```
print(res)
```

```
[[-1. ]
 [ 0.5]]
```

Рис. 10.

4 Проекц матриц (The projection Matrix) P

Өмнөх хэсэгт дурдсан тэгшитгэл (3.9)-ийг À матрицаар хоёр талаас нь үйлчилбэл,

$$A\bar{x} = A(A^T A)^{-1} A^T \vec{b}, \tag{4.1}$$

болох буюу $\vec{p} = A\bar{x}$ гэж тэмдэглэн үүнийг $\vec{b}$-ийн проекц гэж хэлдэг. Тэгвэл доорх хэлбэрт орно

$$\vec{p} = A(A^T A)^{-1} A^T \vec{b}. \tag{4.2}$$

Энэхүү тэгшитгэлийг матриц хэлбэрт бичвэл,

$$P = A(A^T A)^{-1} A^T, \tag{4.3}$$

төдийгүй энэхүү тэгшитгэлийг проекц матриц гэж нэрлэдэг. Иймээс $\vec{p} = P\vec{b}$ вектор нь A-ийн баганан огторгуй бөгөөд $\vec{p} - P\vec{b}$ ортогонал байгуулагч бүрдүүлэгч. Матриц тэгшитгэл (4.3) нь доорх хоёр шинж чанарыг хангадаг:

- Идэмпотент (Idempotent) $P^2 = P$
- Симметрик (Symmetric) $P = P^T$.

4.1 Проекц болон хамгийн бага квадратын арга

Матриц A нь $M \times N$ бөгөөд баганын хувьд ортонормал шинж чанартай,

$$\left(V_i^T V_j = \begin{Bmatrix} 0, & i \neq j \\ 1, & i = j \end{Bmatrix} \right). \tag{4.4}$$

Энд $\bar{x} = A(A^T A)^{-1} A^T \vec{b}$ нь хамгийн бага квадрат шийд бөгөөд $P = A(A^T A)^{-1} A^T$ проекц матриц. Цаашилбал тэгшитгэл (4)-ийн дагуу $(A^T A)$ нь нэгж матриц бөгөөд $P = AA^T$, $\bar{x} = AA^T \vec{b}$ болж хялбарчилагдана. Үүнтэй холбоотой нэгэн жишээ авч үзэе; $y_1 = -1, \quad y_2 = 0, \quad y_3 = 1.$

$$\begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix}. \tag{4.5}$$

энд $\bar{y} = (y_1 + y_2 + \ldots + y_N)/N$ хувиргалтыг хийж болно. Иймээс бид тэгшитгэл $z = C + Dy$-тэй ажиллахын оронд $z = c + d(y - \bar{y})$-тэй ажиллаж болно,

$$\begin{pmatrix} 1 & y_1 - \bar{y} \\ 1 & y_2 - \bar{y} \\ 1 & y_3 - \bar{y} \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix}. \tag{4.6}$$

Хэрэв энэ тэгшитгэлийг тэгшитгэл (3.13)-тэй ижил хэлбэрт оруулбал, бид шийдийг маш хялбархан олж болно;

$$c = \frac{\vec{a}_1^T \vec{z}}{\vec{a}_1^T \vec{a}_1} = \frac{\sum z_i}{N}, \quad d = \frac{\vec{a}_2^T \vec{z}}{\vec{a}_2^T \vec{a}_2} = \frac{\sum_i (y_i - \bar{y}) z_i}{\sum_i (y_i - \bar{y})^2}. \tag{4.7}$$

Энд $\vec{a}_1, \vec{a}_2$ нь матриц А-ийн баганын элементүүд $A = [\vec{a}_1, \vec{a}_2]$ юм.

PART

III

Хүн-Машин

> Remark (Л. Түдэв, Оройгүй Сүм). Тэнгэр баганагүй нэг дутуу
> Уул бүслүүргүй нэг дутуу
> Далай бүлүүргүй нэг дутуу

5 Би-сплайн арга

Би-сплайнууд нь полиномиал d зэргийн суурь функц бөгөөд дурын сплайн функцыг үүсгэх боломжтой. d полиномиал зэргийн M ширхэг Би-сплайн функцүүдийг өгөгдсөн цэгэн векторын (knot vector) хувьд үүсгэх боломжтой. Цэгэн векторийн элементүүд нь t_j $(j = 1 \ldots M + d + 1)$. Цэгэн векторууд нь эрэмблэгдсэн байх ёстой $(t_1 \leq t_2 \leq \cdots \leq t_{M+d+1})$ бөгөөд хөрөөний ирэн полиномиал функцын завсрыг тодорхойлдог. d зэргийн хөрөөний ирэн полиномиал функцээс бүрдэх дурын полиномиал функц $S_{d,t}$ -ийг доорх байдлаар илэрхийлэх боломжтой

$$S_{d,t}(x) = \sum_{i=1}^{M} \alpha_i B_{i,d,t}(x). \tag{5.1}$$

Энд α_i нь бидний тодорхойлох хэрэгтэй чөлөөт параметр юм. Үр дүн функц нь өөр өөр интервалд тодорхойлогдсон $(t_i \leq x \leq t_{i+1})$ d зэргийн хөрөөний ирэн полиномиудын цуглуулагаар тодорхойлогдоно. Би-сплайн функц өөрөө хөрөөний ирэн полиномуудын цуглуулагааар тодорхойлогддог бөгөөд доорх рекурент шинж чанарыг хангадаг

$$B_{i,0,t}(x) = \begin{Bmatrix} 1 \; if & t_i \leq x \leq t_{i+1}, \\ 0 & otherwise, \end{Bmatrix} \tag{5.2}$$

Энэ нөхцлийг хангаж буй ерөнхий тэгшитгэл нь

$$B_{i,d,t}(x) = \frac{x - t_i}{t_{i+d} - t_i} B_{i,d-1,t}(x) + \frac{t_{i+d+1} - x}{t_{i+d+1} - t_{i+1}} B_{i+1,d-1,t}(x). \tag{5.3}$$

Цэгэн векторыг зураг тус бүрд өөр өөр өгсөн байгааг дээрх зурагнаас харах боломжтой. Рис. 10 ийн хувьд $\{0, 0, 0.33, 0.67, 1, 1\}$, мөн Рис. 11-ийн хувьд цэгэн вектор нь $\{0, 0, 0, 0.11, 0.22, 0.33, 0.44, 0.56, 0.67, 0.78, 0.89, 1, 1, 1\}$. Цэгэн векторууд нь давтагдах боломжтой. Хэрэв цэгэн вектрын эхний $d + 1$ болон сүүлийн $d+1$ элементүүд нь тэнцүү байвал цэгэн векторуудыг БЭХЛЭГДСЭН (clamped) гэж хэлдэг.

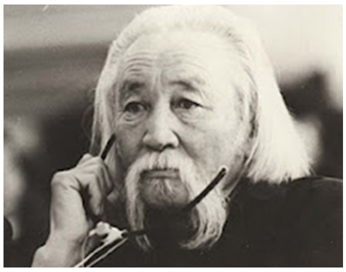


Рис. 11. "Еншөөбү овогт Бямбын Ринчен(1905.11.21—1977.03.04) Монголын орчин үеийн утга зохиолыг үндэслэгчдийн нэг, XX зууны манлай монголч эрдэмтэн байв. Англи, франц, герман, чех, польш, эсперанто, орос хэлийг гаргууд эзэмшсэн бөгөөд эсперанто хэлний өөрөө сурах бичиг зохиосон. Тэрээр тухайн үеийнхээ БНМАУ болон ЗХУ-ын төрийн тэргүүнүүдэд өөрийг нь гүтгэн доромжилсныг эсэргүүцэж, мөн нийгэмд байгаа гажуудлын талаар захидал бичин цаг үеэ шүүмжилж байв. Монголын шинжлэх ухааны зүтгэлтэн, зохиолч, орчуулагч, соён гэгээрүүлэгч. 1961 онд БНМАУ-ын ШУА-ийн жинхэнэ гишүүн, академичаар сонгогдсон."

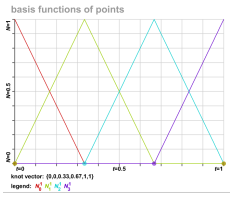


Рис. 12. Би-сплайн суурь функцыг $d = 1$.

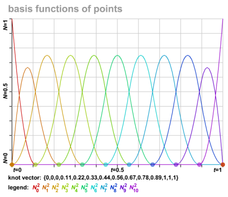


Рис. 13. Би-сплайн суурь функцыг $d = 2$.

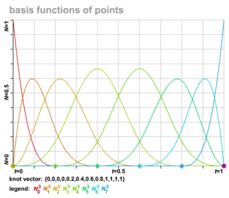


Рис. 14. Би-сплайн суурь функцыг $d = 3$.

$$\underbrace{t_1, \dots t_{d+1}}_{\text{d+1 тэнцүү}}, \underbrace{t_{d+2} \dots t_M}_{\text{M-d-1 дотоод}}, \underbrace{t_{M+1} \dots t_{M+d+1}}_{\text{d+1 тэнцүү элементүүд}}$$

БЭХЛЭГДСЭН (clamped) элементүүдтэй Би-сплайн полиномын хувьд эхний болон сүүлийн Би-сплайн нь нэгтэй тэнцүү ($B_{1,d,t}(t_1) = B_{M,d,t}(t_{M+d+1}) = 1$). Харин бусад элементүүдийн хувьд тэгтэй тэнцүү ($B_{i,d,t}(t_1) = B_{i,d,t}(t_{M+d+1}) = 0, \quad i = 2 \dots M-1$). Дотоод цэгэн элементүүд $M-d-1$ нь өөр хоорондоо ижил биш үед үр дүн сплайн функц нь $d-1$ зэрэг хүртэл тасралтгүй уламжлалтай. Иймээс дурын полиномыг доорх байдлаар тодорхойлж болно,

$$S_{d,t}(x) = \sum_{i=1}^{M} \alpha_i B_{i,d,t}(x), \quad t_{d+1} \leq x < t_{M+1}. \tag{5.4}$$

N дата $\{x_i, y_i\}, i = 1, \dots, N$ нөхцлийг ($M \leq N$) хангах үед, бид тэгшитгэл (5.4)-ийн параметр α_i хамгийн бага квадратын аргатай ижил замаар олох боломжтой. Хэрэв $M = N$ бол, полином $S_{d,t}(x)$ нь бүх датаг дайрч байхаар байгуулах боломжтой. Анхааруулж хэлэхэд нэг асуудал нь датаны нэг утгыг нь өөрчилөхөд, бүх параметрүүдийг дахин тооцоолох хэрэгтэй болдог. Энэ асуудалыг шийдэхийн тулд өөр нэгэн аргыг авч үзэе. Тэрхүү аргыг NURBS гэж нэрлэдэг.

6 NURBS

Муруй шугаман датаг ойролцоолох функцыг олох нь төвөгтэй бөгөөд төгс дөхүүлэн, ойролцоолох $f(x) \approx y(x)$ бараг боломжгүй юм. Энэ тохиолдолд датаг параметрээс $x(s), y(s)$ хамааруулан тодорхойлсоны дараа, дурын полиномиал функцээр ойролцоолно. Энэхүү ойролцоолж буй аргыг NURBS гэж хэлдэг. NURBS нь датаг параметрт тэгшитгэлд хувилгах эсвэл жин харгалзуулан ойролцоолох юм. Ойролцоолж буй үр дүн функц нь дата цэгийг дайрж гарахгүй бөгөөд тасралтгүй уламжлалтай.

Өгөгдсөн N дата $\vec{p} = \{x_i, y_i\}, i = 1, \ldots, N$. Бид одоо d зэргийн, $t_j, j = 1, \ldots, N+d+1$ цэгэн вектортой N Би-сплайн $B_{i,d,t}(s)$-ийг тодорхойлно. Цаашилбал бид дата бүрд жин w_i мөн датаг Би-сплайн ашиглан ойролцоолно. Үүний дараа муруй шугамыг параметр тэгшитгэлд шилжүүлэнэ $\vec{p}(s) = \{x_i(s), y_i(s)\}, t_{d+1} \leq s \leq t_{N+1}$ доорх аргаар

$$\vec{p}(s) = \sum_{i=1}^{N} R_{i,d,t} \vec{p}_i, \quad R_{i,d,t} = \frac{B_{i,d,t}(s) w_i}{\sum_j B_{j,d,t}(s) w_j}. \tag{6.1}$$

Энд Би-сплайныг $B_{i,d,t}(s)$ тэгшитгэл (5.2) ба (5.3)-д тодорхойлсон. NURBS нь Би-сплайн болон Bezier муруйн ерөнхий хэлбэр юм.

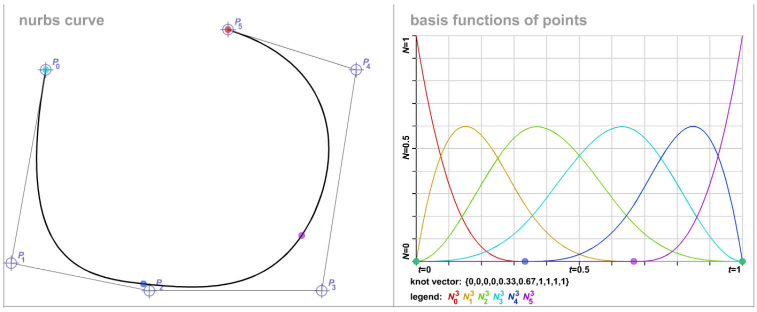


Рис. 15. 6-н цэг $P_0, \ldots, P_5$ өгөгдсөн үед кубик полином хэрхэн ойролцоолж буйг үзүүлэв. Эх сурвалж: http://geometrie.foretnik.net/files/NURBS-en.swf

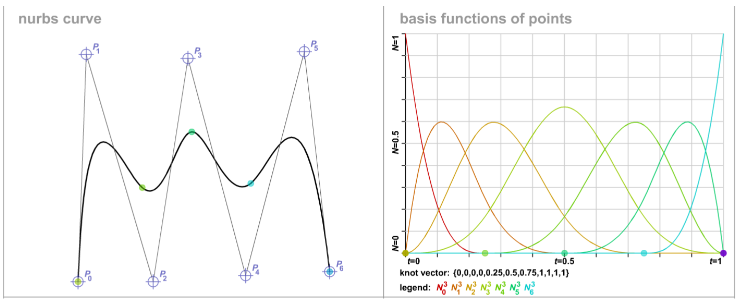


Рис. 16. 6-н цэг $P_0, \ldots, P_6$ өгөгдсөн үед кубик полином хэрхэн ойролцоолж буйг үзүүлэв.

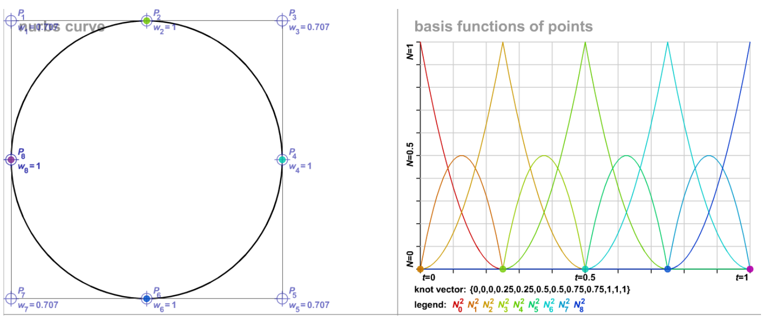


Рис. 17. 9-н цэг $P_0, \ldots, P_9$ мөн цэг бүрд жин w_i өгөгдсөн үед кубик полином хэрхэн ойролцоолж буйг үзүүлэв. Жин нь цэг бүрд ижил буюу $w_i = 1$ бол Bezier муруй болно.

PART

IV

Хүн-Машин

Remark (Л. Түдэв, Оройгүй Сүм). Өндөр тэнгэр үүлэн давхаргатай
Өвсөн дэлхий цасан давхаргатай
Өргөн далай мөсөн давхаргатай

7 Ортогонал функц

Полиномиал ойролцоололын хувьд өндөр эрэмбийн зэргийн коэффицентүүдийг, $f(x) \approx \alpha_3 x^3$, нь тооцохгүй, зөвхөн бага зэргийн коэффицентүүдийг (α_1, α_2) нь функц $f(x) \approx \alpha_1 x + \alpha_2 x^2$-ээс тооцож олсон гэж үзэе. Хэрэв бага эрэмбийн коэффицентүүдэд өндөр эрэмбийн коэффицентүүдийг нэмж, ойролцоолж буй функцын зэргийн эрэмбийг өөрчилөхөд, $f(x) \approx \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3$, бага эрэмбийн коэффицентүүдийг (α_1, α_2) дахин тооцох шаардлагатай болдог. Маш олон хэрэглээнд, бага эрэмбийн тооцож олсон параметрүүдийн утганд өндөр эрэмбийн параметрүүдийг нэмхэд бага эрэмбийн параметрийг дахин тооцохгүй байх нь чухал байдаг. Өөр үгээр хэлбэл тооцож олсон утгыг дахин тооцож олохгүй байхаар функцын эрэмбийг ихэсгэх нь чухал байдаг.

Бага эрэмбийн функц дээр нэмэлт дээд эрэмбийн хэсгийг нэмж нарийвчилан тооцоход бага эрэмбийн параметрийг дахин тооцохгүй байхад хэрэг болдог функц бол Ортогонал функц юм. Ортогонал функцүүд нь ортогонал хувиргалттай ижил шинж чанартай юм. $y(x)$-ийн ойролцоололыг суурь функц $\phi_i(x)$-ээр илэрхийлбэл

$$y(x) = \sum_{i=1}^{M} \alpha_i(x) \phi_i(x), \tag{7.1}$$

суурь функц нь ортогонал шинж чанарыг хангадаг

$$\langle \phi_i(x) \phi_j(x) \rangle = \int_{-\infty}^{\infty} \phi_i(x) \phi_j(x) dx = \delta_{ij}, \tag{7.2}$$

Энд

$$\delta_{ij} = \begin{Bmatrix} 1 & i = j \\ 0 & i \neq j \end{Bmatrix} \tag{7.3}$$

шинж чанарыг хангадаг. Энэ шинж чанарыг нь ашиглан бид тэгшитгэл (7.1)-ийн параметрийг тодорхойлж болно,

Рис. 18. “Чадраабалын Лодойдамба (1917 оны 8 сарын 20-нд Говь-Алтай аймагт; † 1970 оны 1 сарын 11-нд) нь БНМАУ-ын Төрийн хошой шагналт (1954, 1971) Монголын зохиолч байв. 1940 оноос Холбооны техникумд физикийн багш, 1941 онд Радио хороонд орлогч дарга, 1944-45 онд МУИС-ийн Номын төв сургуульд түүх, гүн ухааны багш, МЗЭ-ийн хорооны Цог сэтгүүлийн эрхлэгч, 1959 оноос Урлагийн хэрэг эрхлэх хорооны 1-р орлогч дарга, Соёлын яамны 1-р орлогч сайд, Уран сайхны зөвлөлийн даргаар ажиллаж байсан.”

$$
\begin{aligned}
\int_{-\infty}^{\infty} y(x)\phi_j(x)dx &= \int_{-\infty}^{\infty} \phi_j(x)\sum_{i=1}^{M}\alpha_i(x)\phi_i(x)dx \\
&= \sum_{i=1}^{M}\alpha_i\int_{-\infty}^{\infty}\phi_j(x)\phi_i(x)dx \\
&= \sum_{i=1}^{M}\alpha_i\delta_{ij} = \alpha_j.
\end{aligned} \tag{7.4}
$$

Иймээс параметрийг олохын тулд

$$
\alpha_j = \int_{-\infty}^{\infty} y(x)\phi_j(x)dx, \tag{7.5}
$$

тэгшитгэлийг ашиглана. Туршилтын үр дүн $\{x_i, y_i\}$ мэдэгдэж байгаа, харин функцын хамаарлын тэгшитгэл $y(x)$ мэдэгдэхгүй байгаа тохиолдолыг сонирхоё. Энэ тохиолдолд тэгшитгэл (7.5) интегралыг бодож параметрийг олох боломжгүй юм. хэдий тиймч суурь функц ортонормал шинж чанарыг хангаж байх ёстой бөгөөд хэмжилтийн магадлалын нягтын түгэлт нь ортогонал шинж чанарыг нь хангаж байх ёстой,

$$
\langle\phi_i(x)\phi_j(x)\rangle = \int_{-\infty}^{\infty}\phi_i(x)\phi_j(x)p(x)dx = \delta_{ij}, \tag{7.6}
$$

бөгөөд параметр нь

$$
\alpha_j = \int_{-\infty}^{\infty} y(x)\phi_j(x)p(x)dx \tag{7.7}
$$

гэж тодорхойлогдоно. Магадлалын нягтын функцыг нь доорх байдлаар тодорхойлоё

$$
p(x) \approx \frac{1}{N}\sum_{n=1}^{N}\delta(x - x_n). \tag{7.8}
$$

Делта (Дирак делта) функц нь

$$
\int_{-\infty}^{\infty} f(x)\delta(x - x_n) = f(x_n) \tag{7.9}
$$

тодорхойлогддог. Иймээс параметрийг олохын тулд

$$
\alpha_i \approx \frac{1}{N}\sum_{n=1}^{N} y_n\phi_i(x_n), \tag{7.10}
$$

ашиглана.

8 Хэрхэн ортонормал суурь функц үүсгэх вэ?

Хэрэв бидэнд ортонормал шинж чанарыг (тэгшитгэл (7.6)) хангадаггүй $g_i(x)$ суурь функцүүд өгөгдсөн бол Грамм Шмидтийн ортогоналчилах аргыг ашиглан ортонормал суурь функцүүдийг $\phi_i(x)$ үүсгэх боломжтой;

$$\begin{aligned}
\phi_1(x) &= \frac{g_1(x)}{(\int g_1(x)g_1(x)dx)^{1/2}},\\
\tilde{\phi}_2(x) &= g_2(x) - \phi_1(x)\int \phi_1(x)g_2(x)dx,\\
\tilde{\phi}_3(x) &= g_3 - \phi_1(x)\int \phi_2(x)g_3(x)dx - \phi_2(x)\int \phi_2(x)g_3(x)dx,\\
&\vdots\\
\tilde{\phi}_n(x) &= g_n - \sum_{i=1}^{N-1}\phi_i(x)\int \phi_i(x)g_n(x)dx,\\
\phi_i(x) &= \frac{\tilde{\phi}_i(x)}{(\int \tilde{\phi}_i(x)\tilde{\phi}_i(x)dx)^{1/2}}.
\end{aligned} \tag{8.1}$$

Энэхүү ортогонал суурь функц үүсгэх арга нь дурын вектор суурьнаас ортогонал суурь вектор үүсгэхтэй төстэй юм. Бидэнд дата өгөгдсөн үед ортогонал функц нь хэмжилтийн таамагласан утгаар (expected value) тодорхойлогддог,

$$\begin{aligned}
\tilde{\phi}_k(x) &= g_n - \sum_{i=1}^{N-1}\phi_i(x)\int \phi_i(x)g_k(x)p(x)dx,\\
\phi_i(x) &= \frac{\tilde{\phi}_i(x)}{(\int \tilde{\phi}_i(x)\tilde{\phi}_i(x)p(x)dx)^{1/2}},
\end{aligned} \tag{8.2}$$

энд $p(x) \approx \exp\{-x^2\}$ бол Хермитийн полином, $p(x) \approx \exp\{-x\}$ бол Лагуэрийн полиномууд, харин $p(x) \approx (1-x^2)^{-1/2}$ Чебишевийн полином гэж нэрлэдэг. Хэрэв N дата өгөгдсөн бол интегралыг нийлбэрээр солих хэрэгтэй буюу

$$\begin{aligned}
\tilde{\phi}_k(x) &= g_n - \sum_{i=1}^{N-1}\phi_i(x)\sum_{n=1}^{N}\phi_i(x_n)g_k(x_n)p(x_n),\\
\phi_i(x) &= \frac{\tilde{\phi}_i(x)}{(\sum_1^N \tilde{\phi}_i(x_n)\tilde{\phi}_i(x_n)p(x_n))^{1/2}}.
\end{aligned} \tag{8.3}$$

8.1 Хэрэглээ

Бидэнд суурь функц $g_1 = 1, g_2 = x, g_3 = x_2$ өгөгдсөн, математикт $g = \{1, x, x^2\}$ гэж тэмдэглэдэг, бөгөөд хэмжилтийн утгууд { (-3,0), (-1,1/4), (1, 1/2), (3,1)} нь гэж тодорхойлогдмон. Тэгвэл өгөгдөлд зориулж ашиглан ортогонал функцыг

$\tilde{\phi} = \{\tilde{\phi}_1, \tilde{\phi}_2, \tilde{\phi}_3\}$ Грам Шмидтийн арга ашиглан байгуулая. Эхний алхам,

$$\begin{aligned}
\tilde{\phi}_1(x) &= \frac{g_1(x)}{(\sum_{n=1}^{N} g_1(x_n)g_1(x_n)p(x_n))^{1/2}} = \frac{\tilde{\tilde{\phi}}_1(x)}{\langle\tilde{\tilde{\phi}}_1(x)\tilde{\tilde{\phi}}_1(x)\rangle^{(1/2)}},\\
\tilde{\phi}_1(x) &= \frac{1}{(g_1(x_1)g_1(x_1)p(x_1) + g_1(x_2)g_1(x_2)p(x_2) + g_1(x_3)g_1(x_3)p(x_3))^{1/2}},\\
\tilde{\phi}_1(x) &= 1,
\end{aligned} \tag{8.4}$$

энд $g_1 = g_1(x_1) = g_1(x_2) = g_1(x_3) = 1$. Хоёр дугаар алхам $\tilde{\phi}_2(x)$,

$$\begin{aligned}
\tilde{\tilde{\phi}}_2(x) &= g_2(x) - \phi_1(x)\frac{1}{N}\sum_{n=1}^{N}\phi_1(x_n)g_2(x_n),\\
\tilde{\tilde{\phi}}_2(x) &= x - \phi_1(x)\frac{1}{4}\sum_{n=1}^{4}\phi_1(x_n)x_n,\\
\tilde{\tilde{\phi}}_2(x) &= x - \phi_1(x)\frac{1}{4}(\phi_1(x_1)x_1 + \phi_1(x_2)x_2 + \phi_1(x_3)x_3 + \phi_1(x_4)x_4),\\
\tilde{\tilde{\phi}}_2(x) &= x - 1\frac{1}{4}(1(-3) + 1(-1) + 1(1) + 1(3)) = x\\
\tilde{\phi}_2(x) &= \frac{\tilde{\tilde{\phi}}_2(x)}{\langle\tilde{\tilde{\phi}}_2(x)\tilde{\tilde{\phi}}_2(x)\rangle^{(1/2)}} = \frac{x}{\sqrt{5}}.
\end{aligned} \tag{8.5}$$

гэх мэтчилэн тэгшитгэл (8.3)-ийг ашиглан хамгийн сүүлийн ортогонал суурь $\phi_3(x)$,

$$\begin{aligned}
\tilde{\tilde{\phi}}_3(x) &= g_3(x) - \phi_1(x)\sum_{i=1}^{4}\phi_2(x_n)g_3(x_n) - \phi_2(x)\sum_{i=1}^{4}\phi_2(x_n)g_3(x_n),\\
\tilde{\tilde{\phi}}_3(x) &= g_3(x) - \phi_1(x)\sum_{i=1}^{4}\phi_2(x_n)g_3(x_n) - \phi_2(x)\sum_{i=1}^{4}\phi_2(x_n)g_3(x_n),\\
\tilde{\tilde{\phi}}_3(x) &= x^2 - 1\sum_{i=1}^{4}\phi_2(x_n)g_3(x_n) - \sum_{i=1}^{4}\phi_2(x_n)g_3(x_n) = x^2 - 5,\\
\tilde{\phi}_3(x) &= \frac{\tilde{\tilde{\phi}}_3(x)}{\langle\tilde{\tilde{\phi}}_3(x)\tilde{\tilde{\phi}}_3(x)\rangle^{(1/2)}} = \frac{x^2-5}{4}.
\end{aligned} \tag{8.6}$$

Иймээс бидний ортогонал суурь функц нь $\tilde{\phi} = \{\tilde{\phi}_1, \tilde{\phi}_2, \tilde{\phi}_3\} = \{1, \frac{x}{\sqrt{5}}, \frac{x^2-5}{4}\}$. Түүнчилэн бидэнд шугаман, нэгдүгээр эрэмбийн тэгшитгэл өгөгдсөн гэж үзэе. Дээр олсон ортогонал суурь функц болон тэгшитгэл (7.1), (7.10)-ийг ашиглан

(??)

$$\begin{aligned}
f(x) &= \alpha_1\tilde{\phi}_1 + \alpha_2\tilde{\phi}_2,\\
\alpha_1 &= \frac{7}{16}, \alpha_2 = \frac{13}{16\sqrt{5}}.
\end{aligned} \tag{8.7}$$

Одоо ортогонал функцын шинж чанар болох өндөр эрэмбийн коэффицентийг нэмхэд бага эрэмбийн коэффицентүүдийг дахин тооцох шаардлагагүй болдог буюу тэгшитгэл (8.4) нь квадрат эрэмбийн хувьд

(??)
$$\begin{aligned} f(x) &= \alpha_1\tilde{\phi}_1 + \alpha_2\tilde{\phi}_2 + \alpha_3\tilde{\phi}_3, \\ \alpha_3 &= \frac{1}{16}, \end{aligned} \tag{8.8}$$

энд ортогонал функцын дагуу коэффицентүүд (α_1, α_2) нь тэгшитгэл (??)-д олсонтой ижил утгатай учир дахин тооцох шаардлагагүй.

Хүн-Машин

PART

V

> Remark (Л. Түдэв, Оройгүй Сүм). Алгандаа хөхтэй байдаг арслан нэг хачин
> Амаа байтал шилээрээ дуугардаг голио нэг хачин
> Ам нь хэнхдэгэндээ байдаг заан нэг хачин

9 Ортогонал матриц

Ортогонал матриц нь ортонормал баганатай матриц юм. Ортогонал матриц нь доорх шинж чанартай,

$$Q^TQ = I = QQ^T, \quad Q^T = Q^{-1}; \tag{9.1}$$

$$||Q\vec{x}|| = \vec{x}, \quad (Q\vec{x})^T(Q\vec{y}) = \vec{x}^T\vec{y}. \tag{9.2}$$

Баганууд нь шугаман хамааралгүй дурын матриц A-ийг дараах байдлаар задлаж болно,

$$A = QR, \tag{9.3}$$

энд Q-ийн баганууд нь ортонормал бөгөөд R нь дээд гурвалжин (upper triangular) ба урвуу (invertible) нь оршин байдаг. Энэ шинж чанарыг нь ашиглан, шугаман тэгшитгэл $A\vec{x} = \vec{b}$-ийн шийдийг олох боломжтой буюу доорх хэлбэрт орно,

$$\bar{x} = R^{-1}Q^T\vec{b}. \tag{9.4}$$

9.1 Псюдо-урвуу (Pseudo inverse)

Шийд $\bar{x}$-ийг хэрхэн ойролцоогоор олох аргыг дурын матриц A болон вектор $\vec{b}$-ын хувьд боловсруулах. Оптимал сонголт нь проекц вектор $\vec{p} = A\bar{x}$ нь $\vec{b}$-д хамгийн ойрхон байх юм. Хамгийн ойрхон проекц векторыг олохын тулд доорх тэгшитгэлийн хамгийн бага шийдийг олох хэрэгтэй,

$$E = ||A\vec{x} - \vec{b}||. \tag{9.5}$$

Өмнө харсанчлан хэрэв A матриц нь кватрат бөгөөд урвуу нь оршин байдаг бол шугаман тэгшитгэлийн шийд нь $\bar{x} = A^{-1}\vec{b}$. Харин өгөгдсөн $M \times N$ багана ба мөртэй A матрицын хувьд парамерийн тоо нь тэгшитгэлийн тооноос бага үед $N > M$ хамгийн бага квадратын аргыг ашиглана. Нөгөө талаасаа, $N < M$ үед $A\bar{x} = \vec{p} = P\vec{b}$ шинжийг хангасан маш олон шийд оршин байдаг.

Рис. 19. “Ренчиний Чойном нь Монголын алдартай яруу найрагч байсан. Тэрээр Хэнтий аймгийн Дархан сумaнд 1936 онд төржээ. "Гал морин цаг "Залуу нас "Сүмтэй бударын чулуу "Тал "Улаан дэвтэр "Хүн"зэрэг бүтээл туурвижээ. 1990 онд Монгол Улсын төрийн шагналыг нэхэн олгожээ.”

СОНГОЛТ: Хамгийн оптимал шийд нь $A\bar{x} = \vec{p}$-ны уртын хэмжээ хамгийн богинотой нь юм. Тэгшитгэлийн $A\vec{x} = \vec{b}$ шийдийг нь $\bar{x} = A^{+}\vec{b}$ хэлбэрт бичих буюу A^{+}-г псюдо-урвуу (pseudo inverse) гэж хэлдэг.

Жишээ нь: Өгөгдсөн нь,

$$A = \begin{pmatrix} \mu_1 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mu_1 > 0, \quad \mu_2 > 0. \tag{9.6}$$

Вектор $\vec{b} = [b_1, b_2, b_3]^T$-ын проекц нь $\vec{p} = P\vec{b} = [b_1, b_2, 0]^T$ болно. Иймээс $A\bar{x} = \vec{p}$,

$$\begin{pmatrix} \mu_1 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}. \tag{9.7}$$

Эндээс бид $\bar{x}_1\mu_1 = b_1$, $\bar{x}_2\mu_2 = b_2$ гэдгийг харж байна. Эсвэл шийдийг $\bar{x} = A^{+}\vec{b}$ гэж олж болно,

$$A^{+}\vec{b} = \begin{pmatrix} \mu_1^{-1} & 0 & 0 \\ 0 & \mu_2^{-1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}. \tag{9.8}$$

Диагоналын элементүүд нь тэгээс их бөгөөд бусад элементүүд нь тэгтэй тэнцүү матриц нь нэгэн онцгой матрицын бүлд харялагддаг. Доорх байдлаар ерөнхий хэлбрийг нь илэрхийлж болно,

$$\Sigma = \begin{pmatrix} \mu_1 & 0 & 0 & 0 & \dots \\ 0 & \mu_2 & 0 & 0 & \dots \\ 0 & 0 & \mu_3 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \rightarrow \Sigma^{+} = \begin{pmatrix} \mu_1^{-1} & 0 & 0 & 0 & \dots \\ 0 & \mu_2^{-1} & 0 & 0 & \dots \\ 0 & 0 & \mu_3^{-1} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{9.9}$$

10 Онцгой утгын задрал (Singular Value Decomposition)

Дурын $N \times N$ хэмжээт матриц А-г задлан бичих боломжтой,

$$A = Q_1 \Sigma Q_2^T. \tag{10.1}$$

- Q_1 нь $M \times M$ хэмжээст ортогонал матриц.
- Q_2 нь $N \times N$ хэмжээст ортогонал матриц.
- Σ нь $M \times N$ хэмжээст дээр дурдсанчилан, тэгшитгэл (9.9), диагонал матриц.

Диагоналын утгууд μ_i-ыг матриц A-ын онцгой утга гэж хэлдэг бөгөөд

$$A^+ = Q_2 \Sigma^+ Q_1^+, \tag{10.2}$$

хэлбэрт оруулж болно. Тэгвэл тэгшитгэлийн $A\vec{x} = \vec{b}$ шийд нь,

$$\bar{x} = Q_2 \Sigma^+ Q_1^+ \vec{b}. \tag{10.3}$$

Оцгой утгын задрал (SVD): Онцгой утгын задралыг параметрын тоо нь тэгшитгэлийн тооноос их $M > N$ (under-determined system) эсвэл $M < N$ (over-determined system) систем тэгшитгэлийн шийдийг олоход хэрэглэх боломжтой.

Жишээ програм: Онцгой Утгын Задрал

```
In [36]: import numpy.linalg as linalg
         import numpy as np
```

Дурын матрицын Онцгой Утгын Задралыг авч үзэе.

```
In [37]: A = np.random.randn(3, 4)
```

Өгөгдсөн матриц нь:

```
In [38]: print(A)

[[-0.79164636  0.00898644  0.66668641 -0.80577924]
 [ 0.18219955  0.04515089 -0.31628269  1.64914704]
 [ 0.62033301 -0.54341104 -0.87410559 -0.67716419]]
```

Тэгвэл матриц $A = Q_1 \Sigma Q_2^T$ нь:

```
In [42]: Q1, Sigma, Q2_T = linalg.svd(A, full_matrices=True)
```

```
In [43]: print(Q1)

[[ 0.54737268  0.4417076  -0.71082877]
 [-0.82363009  0.13373172 -0.55113455]
 [ 0.14837997 -0.88713597 -0.43700476]]
```

```
In [44]: print(Sigma)

[2.02712881 1.51045362 0.29578851]
```

```
In [45]: print(Q2_T)

[[-0.24238501 -0.05569454  0.24454594 -0.93720057]
 [-0.57971332  0.32578754  0.68034789  0.30809365]
 [ 0.64647565  0.69712368  0.27858668 -0.13593115]
 [-0.43272533  0.63622259 -0.63222899 -0.09086318]]
```

Рис. 20. Дурын матрицын Онцгой Утгын Задрал (SVD)

PART

VI

Хүн-Машин

Remark (Л. Түдэв, Оройгүй Сүм). Чулуун дээр ногоо ургуулах нэг бэрх
Усан дээр гал түлэх нэг бэрх
Үхрийн эвэр дээр цас тогтоох нэг бэрх

11 Нейроны сүлжээ

Рис. 22. “Бэгзийн Явуухулан нь XX зууны манлай их яруу найрагч байв. Залуучуудын үнэн, Улаанбаатарын мэдээ сонинд сурвалжлагч, хариуцлагатай эрхлэгч, МЗЭ-ийн хороонд яруу найргийн зөвлөлийн эрхлэгч, нарийн бичгийн дарга, мэргэжлийн зохиолчоор ажиллаж байсан. XX зууны Монголын яруу найрагт иргэний уянгын шинэ чиглэл бий болгосон төлөөлөгч. 1967 онд Би хаана төрөв, Сохор зоосны дууль, Пионер, Түлээчин, Анхны цас зохиолоороо Төрийн шагнал хүртжээ.”

Компьютер ашиглан шинжлэх ухаан болон инженерийн тооцоо хийхийн тулд тооцон бодох математикийн аргыг ашигладаг. Тооцон бодох математикийн хувьд математик тэгшитгэлийг буюу шинжилж буй функцыг ойролцоолон тооцоолохын тулд тасралтгүй хувьсагч/функцыг тасралттай хувьсагч/функц болгон хувиргаж, тэрхүү тасралттай хувьсагч/функцыг ашиглан өгөгдөлийг шинжлэн, судалдаг. Өгөгдөлийн хэмжээ их түүнчлэн олон нуугдмал параметртээс хамаарч буй үед компьютер ашиглан тооцоо гүйцэтгэх шаардлага гардаг. Тиймээс тасралттай функцын тооцоо хийдэг дүрмийг компьютерийн хэл болох програмд бичдэг.

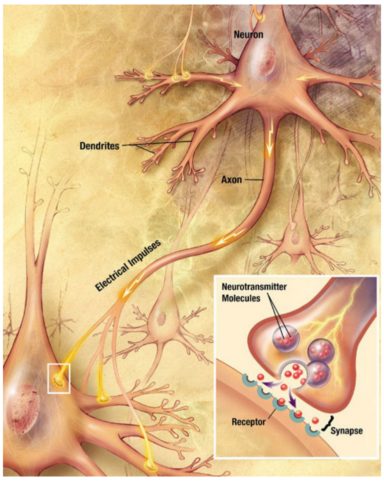


Рис. 21. Нейроны бүтэц

Гэвч бодит амьдралд дата нь гадны шуугиан болон хүрээлэн буй орчны нөлөөлөл ихтэй мөн энгийн нүдээр ялган, ангилхад төвөгтэй байдаг. энэ тохиолдолд компьютерд оруулж өгсөн математик дүрэм маань өгөгдөлд үнэн зөв тооцоо хийх боломжгүй болдог. Энэхүү асуудлыг шийдэхийн тулд судлаачид машин сургалтын алгоритмыг хөгжүүлсэн. Машин сургалтын алгоритм нь компьютерд автоматжуулсан математик дүрэмд захирагддаггүй өгөгдөлд ажиллахад тохиромжтой. Жишээ нь, Машин сургалтын алгоритм нь шугаман биш мөн $n > 3$ хэмжээст огторгуйд харгалзах датаг бүлэглэх, ялгах болон ангилал хийх үйлдэлд маш сайн (жишээ нь: царай таних, обьект илрүүлэх гэх мэт). Машин сургалтын алгоритмын үндсэн философийг нь ойлгохын тулд хүний тархинд мэдээлэл боловсруулах процесс хэрхэн явдагийг сонирхоё. Насанд хүрсэн хүний тархинд гүйцэтгэгдэж байдаг мэдээлэл боловсруулах процесс нь хүүхэд байх үеэс эхэлдэг. Хүүхэд зурагт ном үзэж хүрээлэн буй орчинтойгоо танилцдаг буюу обьектүүдийг хооронд нь ялгаж сурдаг (нохой, муур, хүн, машин гэх мэт). Албан ёсны хэлээр, тэд маш их өгөгдөл цуглуулж, тэрхүү өгөгдөлд суурилж ямар нэгэн оролтын өгөгдөлийг ялган таньдаг. Энэхүү мэдээлэл боловсруулах болон хадгалах процессийг хүний тархи гүйцэтгэдэг. Цаашилбал хүний тархи хэдэн тэр бум нейронуудаас бүрддэг бөгөөд нейрон бүр нь хэдэн мянган бусад нейронуудтай холбогдсон байдаг. Нейрон нь Дендритүүд, сигнал хүлээн авагч, ба нэг аксоноос, сигнал илгээгч, бүрддэг (Рис. 18). Нейроны сүлжээг ашиглан шугаман биш өгөгдөлд ажиллахад ашигладаг бөгөөд математик тооцоололд ашигладаг графикаар илэрхийлбэл: Дендрит нь

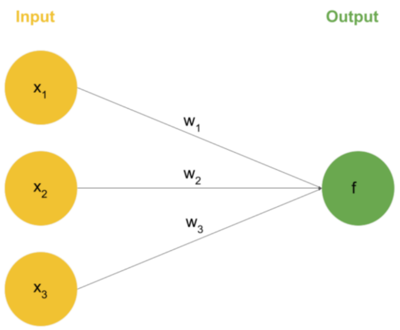


Рис. 23. Нейроны бүтэц: Дендрит ба аксон

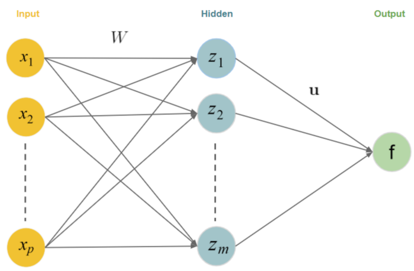


Рис. 24. Нейроны бүтэц: нуугдсан давхаргатай

оролтын датаг x_i илэрхийлдэг бөгөөд аксон нь гаралтын $f(x) = y_j$ утгыг илэрхийлдэг. Оролтын утга бүрд тодорхой жин харгалздаг w_{ij}. Учир нь зарим өгөгдөл бусад өгөгдөлөөсөө илүү чухал байдаг. Жишээ зурагны хувьд $\{i = 1, 2, 3\}$, $j = 1$, мөн $f(x) = y_1$ байна. Тэгвэл гаралтын утгыг $y_1 = \phi(\sum_{i=1}^{3} w_{ij}x_i)$- идэвхижүүлэх функц- гэж тэмдэглэдэг. Нейрон бүр логистик регресстэй ижилхэн таамаглалыг илэрхийлдэг буюу шугаман шийдвэрийн муруйтай. Нейрон бүр гурван алхамыг гүйцэтгэдэг:

- Жинтэй нийлбэрийг тооцох, $\sum_{i=1}^{3} w_{ij}x_i$.
- Идэвхижүүлэх функцыг (activation function) нийлбэрт үйлчилэх, $\phi(\sum_{i=1}^{3} w_{ij}x_i)$.
- Гаралтын нейроны үр дүнг тооцох идэвхижүүлэх функцыг сонгох, $\phi(x)$. Идэвхижүүлэх функцын төрлүүд нь,

Heaviside функц:

$$\phi(x) = \begin{Bmatrix} 1 & x > 0 \\ 0 & else \end{Bmatrix} \tag{11.1}$$

ReLU функц:

$$\phi(x) = ReLU(x) = \begin{Bmatrix} x & x > 0 \\ 0 & else \end{Bmatrix} \tag{11.2}$$

Sigmoid, σ, функц:

$$\phi(x) = \sigma(x) = \frac{\exp(x)}{1 + \exp(x)} \tag{11.3}$$

tanh функц:

$$\phi(x) = \tanh(x). \tag{11.4}$$

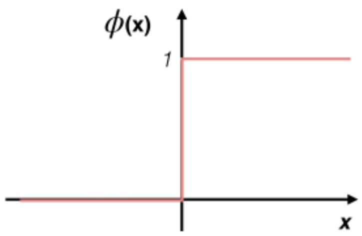


Рис. 25. Heaviside функц.

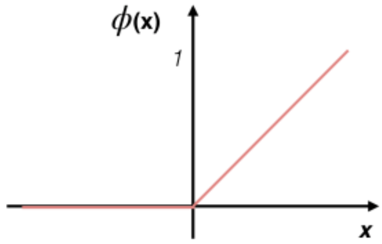


Рис. 26. ReLU функц

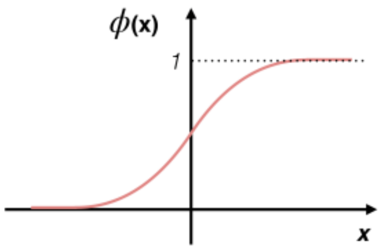


Рис. 27. Sigmoid функц.

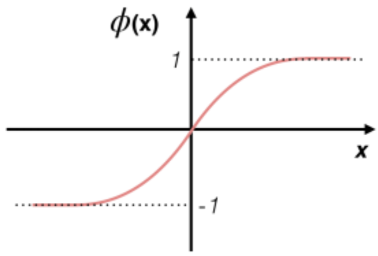


Рис. 28. tanh функц.

Энэхүү зүүнээсээ баруун чиглэлтэй нейроны сүлжээг Урагш Тэжээл Дамжуулалттай Нейроны Сүлжээ (Feedforward Neural Network) гэж нэрлэдэг. Зөвхөн нэг оролт, нэг гаралттай нейроны сүлжээнд нэмэлт нейроны давхаргыг нэмж өгөх боломжтой бөгөөд нуугдсан давхарга гэж хэлдэг. Тэгвэл оролтын нейроны тоог p $(x_1, \ldots, x_p)$ харин нуугдсан давхарга дахь нейроны тоог m $(z_1, \ldots, z_m)$ гэж тэмдэглэе. Оролтын нейрон бүр нь нуугдсан давхаргын нейрон бүртэй шууд холбогддог (Рис. 19). Жишээ нь: оролтын давхаргын нейрон x_1 нуугдсан давхаргын бүх нейронтой $(z_1, \ldots, z_m)$ холбогддог. Цаашилбал нейроны сүлжээ нь дурын тооны $(1, \ldots, L)$ нуугдсан давхаргатай байж болно. Давхарга бүр $m_1, \ldots, m_L$ нейронтой. Зөвхөн нэг нуугдсан давхаргатай үед жин нь (Рис. 19),

$$W = \begin{pmatrix} w_{1,1} & w_{1,2} & \ldots & w_{1,m} \\ w_{2,1} & w_{2,2} & \ldots & w_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ w_{p,1} & w_{p,2} & \ldots & w_{p,m} \end{pmatrix}. \tag{11.5}$$

Хэрэв z_1 нейроны гаралтын утгыг тодорхойлохыг хүсвэл тэгшитгэл (11.5)-ийн нэгдүгээр баганыг сонгон авч,

$$W_1 = \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ \vdots \\ w_{p,1} \end{pmatrix}, \tag{11.6}$$

тэгшитгэл $z_1 = \phi\left(W_1^T x\right)$-ийг тооцно, энд $\phi()$ бол идэвхижүүлэх функц. Цаашилбал нийт гаралтын утгыг,

$$f = \sigma(u^T \phi\left(W^T x\right)), \tag{11.7}$$

σ нь гаралтын утгын идэвхижүүлэх функц, Рис. 19. Нейроны сүлжээ нь l тооны

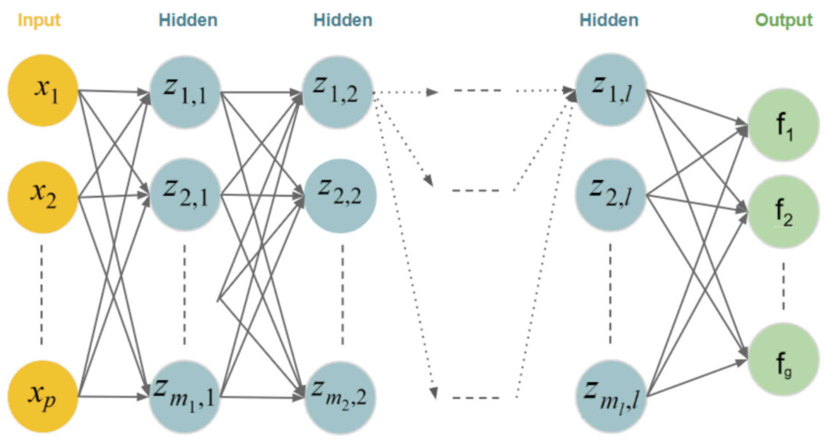


Рис. 29. Нейроны бүтэц

нуугдсан давхаргатай ба g тооны гаралтын утгатай үед дээрх хэлбэрт орно. Жишээ болгон доорх тохиолдолд гаралтын утга нь ямар байхыг тэгшитгэл (11.7)-г ашиглан сонирхоорой.

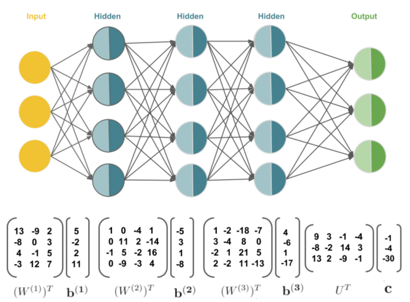


Рис. 30. Нейроны давхарга бүрийн гаралт дах идэвхижүүлэх функц нь $\phi(z) = \max(0, z)$.

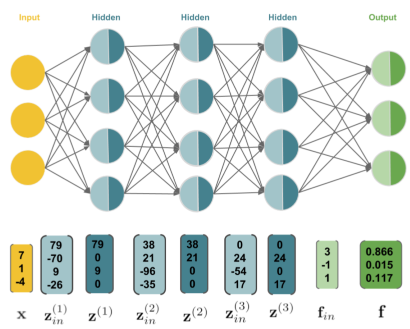


Рис. 31. Нейроны гаралтын идэвхижүүлэх функц нь зигмоид σ.

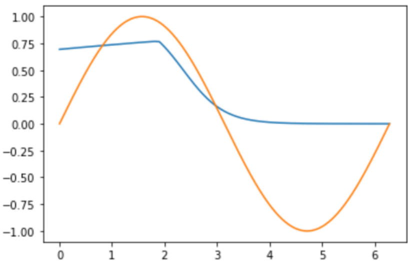


Рис. 32. Sigmoid Идэвхижүүлэх Функцыг гаралтын утганд ашиглаж синус функцыг ойролцоолсон байдал.

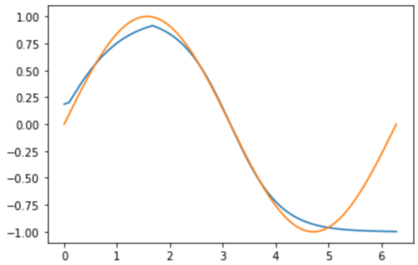


Рис. 33. tanh Идэвхижүүлэх Функцыг гаралтын утганд ашиглаж синус функцыг ойролцоолсон байдал.

11.1 Нейроны сүлжээг сургах

Нейроны сүлжээг сургах хоёр алхам нь:

- Урагш дамжуулах (Forward pass): Мэдээллийн оролт нь нейроны моделоор дамжин таамаглах гаралтын утгыг өгдөг.
- Хойш дамжуулах (Backward Pass): Таамагласан гаралтын утгын алдааны мэдээлэл нь нейрон моделоор хойш урсах бөгөөд нейрон моделийн жинг алдааны утгыг бага байхаар өөрчилдөг.

Алдааны утгыг (Error Value) тооцохын тулд таамагласан нейрон сүлжээний гаралт, $f(x, \theta)$, болон бодит утгын y хоорондох үнийн функцыг (Loss Function), $L(y, f(x, \theta))$, ашигладаг. Хэд хэдэн төрлийн үнийн функц байдаг бөгөөд доор байдлаар ерөнхий тодорхойлж болно,

- Регресс гүйцэтгэх төрлийн өгөгдөлтэй ажиллаж буй үед,

$$L(y, f(x, \theta)) = \frac{1}{2}(y - f(y, \theta))^2.$$

- Ангилах төрлийн өгөгдөлтэй ажиллаж буй үед,

$$L(y, f(x, \theta)) = [y \log(f(x, \theta)) + (1 - y) \log(1 - f(x, \theta))] .$$

Цаашилбал нийт өгөгдөлд тооцсон үнийн функцыг, эрсдэлийн функц (Risk) гэж нэрлэдэг буюу доорх байдлаар илэрхийлдэг,

$$R = \sum_{i=1}^{n} L(y^{(i)}, f(x^{(i)})). \tag{11.8}$$

Нейроны сүлжээ ашиглаж синус функцыг ойролцоосон байдлыг зураг 28, 29-д үзүүлэв.

PART

VII

Хүн-Машин

Remark (Л. Түдэв, Оройгүй Сум). Эрдэнэд алт нэг шар
Эдэнд жанч нэг шар
Идээнд ууранг нэг шар

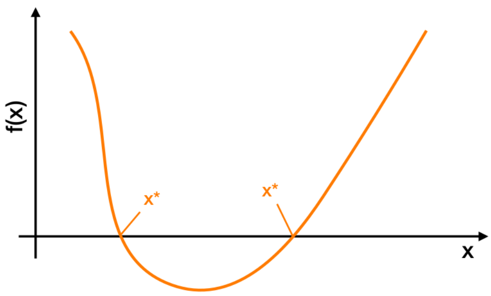
12 Шугаман Бус Систем Тэгшитгэл

Бид бүхэн N хувьсагчтай N шугаман тэгшитгэлийн систем $A\vec{x} = \vec{b}$ -ийн шийдийг хэрхэн олох талаар үзсэн. Энд A нь $N \times N$, харин $\vec{x}, \vec{b}$ нь N хэмжээст. Түүнээс гадна дата өгөгдөлийн хувьд хэрхэн шугаман тэгшитгэлийн шийд $\vec{x}$-ийг олохыг бид үзсэн. Гэвч N шугаман бус тэгшитгэлийн $f_i(\vec{x})$ шийдийг олох нь хялбар биш юм. Шугаман бус систем тэгшитгэлийг $f_i(\vec{x}^*) = 0$ хангах $\vec{x}^*$-ийг $(i = 1, \ldots, N; \vec{x} = (x_1, \ldots, x_N))$ олох нь зорилго юм. Шугамна бус тэгшитгэлийг доорх байдлаар илэрхийлж болно,

$$\begin{aligned} f_1(\vec{x}) &= f_1(x_1, x_2, \ldots x_N) = 0 \\ f_2(\vec{x}) &= f_2(x_1, x_2, \ldots x_N) = 0 \\ &\vdots \\ f_N(\vec{x}) &= f_N(x_1, x_2, \ldots x_N) = 0. \end{aligned} \tag{12.1}$$

Рис. 34. “Хатагин Цэндийн Дамдинсүрэн (1908 оны 9 сарын 14-ий өдөр —1986 оны 5 сарын 27-ий өдөр) БНМАУ-ын Маршал Чойбалсангийн нэрэмжит 3 удаагийн шагналт, анхны Ардын Уран Зохиолч, ШУА-ийн жинхэнэ гишүүн, XX зууны манлай соён гэгээрүүлэгч, орчин цагийн монгол хэлний дүрмийг үндэслэгч эрдэмтэн, орчуулагч, зохиолч.”

Ерөнхий тохиолдолд $\vec{F}(\vec{x}) = 0$ гэж бичдэг. Шугаман бус тэгшитгэлийн системийн шийдийг тооцох нь язгуур олох болдлогын нэгэн төрөл юм. Өөр үгээр илэрхийлбэл $f(x)$ функцын графикийг байгуулан, фукцын x тэнхлэгийг дайрдаг x^* утгыг олох юм.



12.1 Мэдрэмж болон нөхцөл

Хэрэв тэгшитгэлийн шийд оршин байдаг бол, шийдийг нь ямар нэгэн алгоритм хэрэглэн олох боломжтой юу?. Тэгвэл тэгшитгэл $f(x) = 0$-ийг хангадаг ойролцоо шийд нь $\tilde{x}$ харин бодит шийд нь x^* гэж үзэе. Энэ тохиолдолд $\| f(\tilde{x}) \| \approx 0$

буюу $\| \tilde{x} - x^* \| \approx 0$. Бодит шийдийг яг таг олох боломжгүй байдаг бөгөөд харин ойролцоогоор $f(\tilde{x}) \approx 0$-ийн шийд нь $\tilde{x}$ гэж авч үзнэ.

Функц нь Well-Conditioned үед оролтын өгөгдөлийг бага хэмжээгээр өөрчилөхөд үр дүн гаралт мөн бага өөрчлөгддөг. Эсрэг тохиолдол болох ill-conditioned үед оролтын өгөгдөлийг бага өөрчилөхөд гаралтын өгөгдөлийн өөрчилөлт их байдаг.

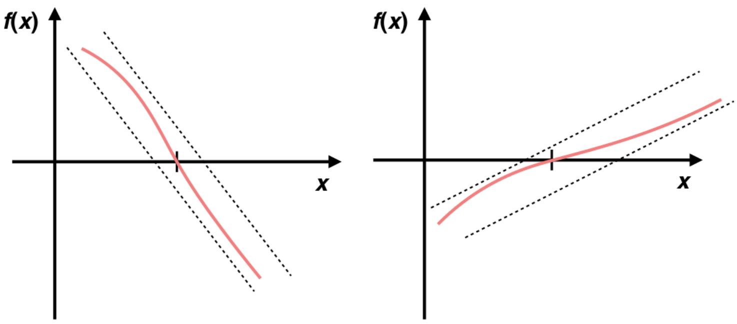
Функцыг ill- эсвэл well-conditioned хэмжихийн тулд нөхцөлийн тоо, k-г тодорхойлдог $k = |\delta y|/|\delta x|$, энд оролтын утгын өөрчилөлт δx, гаралтын үр дүнгийн өөрчилөлт $\delta y = f(x+\delta x) - f(x)$. Хэрэв $f(x+\delta x)$-ийг Тейлорын цуваанд задалбал; $\delta y = (f(x) + f^{'}(x)\delta x) - f(x) = f^{'}(x)\delta x$. нөхцөлийн тоо, k;

$$k = |f^{'}(x)| \tag{12.2}$$

болно. Энэ нь зөвхөн x өгөгдсөн үеп биелдэг бөгөөд бидний тохиолдолд үл мэдэгдэх хувьсагч ба x^* ойролцоо шийд нь мэдэгдэж байгаа. Тиймээс нөхцөлийн тоог доорх байдлаар тодорхойлно,

$$k = \frac{1}{|f^{'}(x^*)|}. \tag{12.3}$$

Хэрэв нөхцөлийн тоо k бага байвал тухайн функцын шийдийг well-conditioned гэж хэлдэг. Энд $|f^{'}(x^*)|$ бага үед функцын шийд нь Well-Conditioned, эсрэг тохиолдолд ill-Conditioned (тангент нь хэвтээ) гэнэ. Графикаар эхнийх нь Well-Conditioned, сүүлийнх ill-Conditioned тохиолдолууд.



12.2 Нийлэлтийн Хурд

Шугаман бус тэгшитгэлийн шийдийг олохын тулд давталтын аргыг ашигладаг. Анхны таамагласан утга $x^{(0)}$ нь боломжит шийдүүдийн дарааллыг үүсгэдэг, $x^{(1)}, x^{(2)}, x^{(3)}, \ldots x^{(k)}, \ldots$; буюу $\{x^{(k)}\}_{k=0}^{\infty}$ гэж тэмдэглэдэг. Нийлэлтийн хурдыг тодорхойлохын тулд алдааны утгыг ашигладаг, $\{x^{(k)}\}^{k\to\infty} \to x^*$ үед

$$\lim_{k\to\infty} \frac{|E^{(k+1)}|}{|E^{(k)}|^r} = C, \quad E^{(k)} = x^{(k)} - x^*. \tag{12.4}$$

r нь нийлэлтийн хурдны эрэмбэ. $r = 1$ шугаман нийлэлтийн хурд, $r = 2$ квадрат нийлэлтийн хурд.

13 Bisection method

$N = 1$ үед $f(x)$-ийн шийд x^* нь $f(x^*) = 0$ хангана гэж үзэе. x^* нь $f(x)$-ийн язгуур юм. $f(x^*) = 0$ хангадаг x^*-ийг олох нэгэн энгийн алгоритм нь BISECTION арга юм. BISECTION арга нь доорх байдлаар тодорхойлогддог: $[a, b]$ завсарт $f(x)$ функцын уламжлал оршин байдаг, тасралтгүй буюу $f(a)f(b) < 0$ хангадаг бол, $[a, b]$ завсарт $f(x^*) = 0$ хангах шийд x^* оршино гэж үзвэл,

- $[a, b]$ завсарын дундаж цэгийг сонгоно: $x_1 = a + (b - a)/2 = (a + b)/2$.
  - Хэрэв $f(x^{(1)}) = 0$ бол, $x^* = x^{(1)}$ төгсөв.
  - Хэрэв $f(x^{(1)}) \neq 0$ бол, $f(x^{(1)})$ нь $f(a)$-тэй ижил тэмдэгтэй эсвэл $f(b)$-тэй.
    - Хэрэв $f(x^{(1)})$ нь $f(a)$-тэй ижил тэмдэгтэй бол, шинэ завсар $[x^{(1)}, b]$ гэж тодорхойлно.
    - Хэрэв $f(x^{(1)})$ нь $f(b)$-тэй ижил тэмдэгтэй бол, шинэ завсар $[a, x^{(1)}]$ гэж тодорхойлно.
  - Шинэ завсарт дахин дээрх тооцоог давтна.

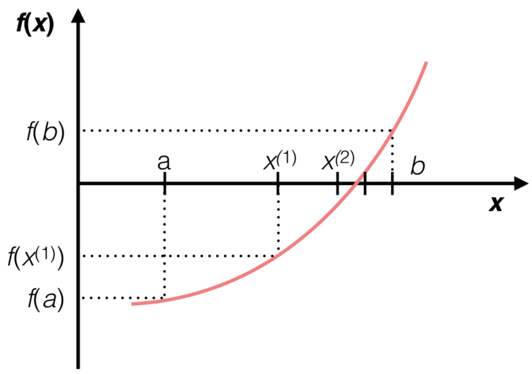


Рис. 35. BISECTION арга.

Алгоритм 1-ийн хувьд N_{max}-ын утгыг зөв тодорхойлох нь чухал юм. Эс бөгөөс бид энэ утгыг таах хэрэгтэй болно. Тиймээс N_{max}-ын утгыг тодорхойлохын тулд бидэнд tol болон $[a, b]$ завсар мэдэгдэж байх ёстой. Bisection аргын хувьд завсар нь үргэлж 1/2 дахин багасч байдаг учир

$$tol = |E^{(k)}| = |x^{(k)} - x^*| \leq (b - a)/2^k \tag{13.1}$$

Algorithm 1 BISECTION АРГА

```
ОРОЛТ: (a, b) завсар, алдааны доод заах tol, хамгийн их давталтын тоо N.
ГАРАЛТ: ойролцоо шийд x^(k) эсвэл шийд олдоогүй nan.
k = 1
FA = f(a)
while k ≤ N do
    x^(k) = (a + b)/2;
    Fx = f(x^(k)).
    if Fx = 0 эсвэл (b − a)/2 < tol then
        Гаралт: x^(k)                ▷ Тэгшитгэлийн шийдийг гаралтанд өгөх.
        Алгоритм Зогсох.
    else
        Гаралт: нийт N давталтын турш шийд олдоогүй nan.
    if sign(FA) = sign(Fx) then
        a = x^(k)
        FA=Fx.
    else
        b = x^(k)                    ▷ Шинэ завсар тодорхойлно.
    k = k + 1
Алгоритм Зогсох.
```

```python
import numpy as np
import math as m
```

```python
def fnc(x):
    return 2*x**3-x**2+x-1
```

```python
N_max=14
tol=10**(-4)
a=0
b=1
k=1
FA=fnc(a)
```

```python
while k<=N_max:
    xk=(a+b)/2
    Fx=fnc(xk)
    if Fx==0 or (b-a)/2<tol:
        print(xk)
        break
    k=k+1
    if FA*Fx>0:
        a=xk
        FA=Fx
    else:
        b=xk
#print('nan')
```

```
0.73895263671875
```

Рис. 36. BISECTION аргад зориулсан Python алгоритм.

нөхцөл хангагдах ёстой. Тэгшитгэл (13.1)-д tol нь доод заах учир k нь хангалттай их буюу $k = N_{max}$ гэж үзвэл,

$$tol = (b-a)/2^k, \quad k = \log_2\left(\frac{b-a}{tol}\right). \tag{13.2}$$

Энд $k = N_{max}$ нь нийт давталтын тоог тодорхойлдог. Дээр жишээ болгон өгсөн Python кодын хувьд N_{max}-ыг олохын тулд тэгшитгэл (13.2) ашиглая. Бидэнд $tol = 10^{-4}$ болон $[0, 1]$ завсар өгөгдсөн. Тэгвэл, $k = N_{max} = \log_2\left(\frac{1}{10^{-4}}\right) = 13.2877 \approx 14$. Энэ үр дүнг ашиглан бид Алгоритм 1-ийн зогсох нөхцлийг нь давталтын тоогоор N_{max} хязгаарлах боломжтой.

Algorithm 2 BISECTION АРГА

ОРОЛТ: (a, b) завсар, алдааны доод зааг tol, хамгийн их давталтын тоо N.
ГАРАЛТ: ойролцоо шийд $x^{(k)}$ эсвэл шийд олдоогүй nan.
$k = 1$
$FA = f(a)$
while $(b - a) > tol$ **and** $k \le N$ **do**
  $x^{(k)} = (a + b)/2;$
  $Fx = f(x^{(k)})$.
  if $sign(FA) = sign(Fx)$ **then**
    $a = x^{(k)}$
    FA=Fx.
  else
    $b = x^{(k)}$ ▷ Шинэ завсар тодорхойлно.
  $k = k + 1$
Алгоритм Зогсох.

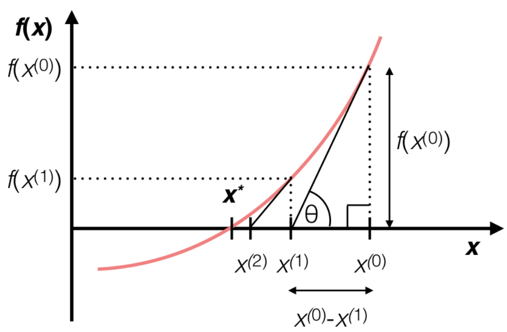
14 Ньютоны Арга

функц $f(x)$-ын уламжлал оршин байдаг бөгөөд уламжлал нь x^* цэгт тэгтэй тэнцүү. Ойролцоо шийд нь $x^{(k)}$, энэ цэгт уламжлал нь тэгтэй тэнцүү биш мөн $|x^* - x^{(k)}|$ маш бага гэж үзэе. Функц $f(x)$-ыг $x^{(k)}$ ойролцоо Тейлорын цуваанд задлан, $x = x^*$ цэгт тооцон, x^*-г олбол:

$$\begin{aligned} f(x^*) &= f(x^{(k)}) + f'(x^{(k)})(x^* - x^{(k)}), \\ 0 &= f(x^{(k)}) + f'(x^{(k)})(x^* - x^{(k)}), \\ x^* &= x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}, \end{aligned} \tag{14.1}$$

энд $f(x^*) = 0$ нөхцлийг ашигласан. Хэрэв эхний таамагласан шийдийн утга нь $x^{(0)}$ бол тэгшитгэл (14.7)-ийн дагуу шийдийн дараалал нь,

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}. \tag{14.2}$$



14.1 Нийлэлтийн хурд

Өмнөх хэсэгт дурдсанчлан уламжлал нь оршин байдаг функц $f(x)$-ын $f(x^*) = 0, f^{'}(x^*) \neq 0$ шийдийг хангадаг $x^* \in [a, b]$ завсарт оршин байдаг. Тэгшитгэл (11.7)-ын Тейлорын цувааны дээд эрэмбийн хэсгийг тооцвол (2-р эрэмбийн уламжлал), бид тэгшитгэл (??)-ийг ашиглан доорх илэрхийлэлийг хялбархан гүйцэтгэх боломжтой,

$$\underbrace{f(x^*)}_{0} = f(x^{(k)}) + f^{'}(x^{(k)})(x^* - x^{(k)}) + \frac{1}{2}f^{''}(x^{(k)})(x^* - x^{(k)})^2, \tag{14.3}$$

$$0 \approx \underbrace{\frac{f(x^{(k)})}{f^{'}(x^{(k)})} - x^{(k)}}_{-x^{(k+1)}} + x^* + \frac{f^{''}(x^{(k)})}{2f^{'}(x^{(k)})}(x^* - x^{(k)})^2, \tag{14.4}$$

$$0 \approx \underbrace{x^* - x^{(k+1)}}_{-E^{(k+1)}} + \frac{f^{''}(x^{(k)})}{2f^{'}(x^{(k)})}\underbrace{(x^* - x^{(k)})^2}_{|E^{(k)}|^2}, \tag{14.5}$$

$$|E^{(k+1)}| = \frac{f^{''}(x^{(k)})}{2f^{'}(x^{(k)})}|E^{(k)}|^2. \tag{14.6}$$

Цаашилбал k нь хангалттай их үед

$$\lim_{k\to\infty}\frac{|E^{(k+1)}|}{|E^{(k)}|^r} = \lim_{k\to\infty}\frac{f^{''}(x^{(k)})}{2f^{'}(x^{(k)})} = \frac{f^{''}(x^*)}{2f^{'}(x^*)} = C. \tag{14.7}$$

Энэ тэгшитгэл нь $r = 2$ үед хангагддаг ба Ньютоны тэгшитгэл нь квадрат хурдаар шийд үрүү дөхдөг. Ньютоны аргын сул тал нь алхам бүрд уламжлалыг нь тооцох хэрэгтэй болдог. Энэхүү хүндрэлийг тойрон гарах аргыг Secant арга гэдэг.

15 Secant Арга

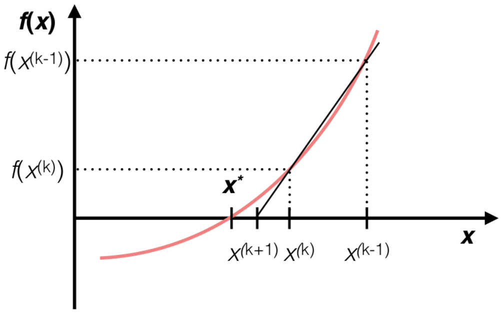
Secant арга нь уламжлалыг тоон аргаар ойролцоолох, буюу

$$f^{'}(x^{(k)}) \approx \frac{f(x^{(k)}) - f(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}. \tag{15.1}$$

тэгвэл энэхүү арга нь доорх байдлаар бичигдэнэ,

$$x^{(k+1)} = x^{(k)} - f(x^{(k)})\frac{x^{(k)} - x^{(k-1)}}{f(x^{(k)}) - f(x^{(k-1)})}. \tag{15.2}$$

Графикаар илэрхийлбэл:


f(x)
f(x(k-1))
f(x(k))
x*
x
x(k+1)
x(k)
x(k-1)

PART

VIII

Хүн-Машин

Remark (Л. Түдэв, Оройгүй Сүм). Оронд орших бэрд хорхой
Биенд орших үхлийн эзэн
Сэтгэлд орших нисваанис

16 Шугаман Бус Систем Тэгшитгэл

Бид бүхэн N хувьсагчтай N шугаман тэгшитгэлийн систем $A\vec{x} = \vec{b}$ -ийн шийдийг хэрхэн олох талаар үзсэн. Энд A нь $N \times N$, харин $\vec{x}, \vec{b}$ нь N хэмжээст. Түүнээс гадна дата өгөгдөлийн хувьд хэрхэн шугаман тэгшитгэлийн шийд $\vec{x}$-ийг олохыг бид үзсэн. Гэвч N шугаман бус тэгшитгэлийн $f_i(\vec{x})$ шийдийг олох нь хялбар биш юм. Шугаман бус систем тэгшитгэлийг $f_i(\vec{x}^*) = 0$ хангах $\vec{x}^*$-ийг $(i = 1, \ldots, N; \vec{x} = (x_1, \ldots, x_N))$ олох нь зорилго юм. Шугамна бус тэгшитгэлийг доорх байдлаар илэрхийлж болно,

$$\begin{aligned} f_1(\vec{x}) &= f_1(x_1, x_2, \ldots x_N) = 0 \\ f_2(\vec{x}) &= f_2(x_1, x_2, \ldots x_N) = 0 \\ &\vdots \\ f_N(\vec{x}) &= f_N(x_1, x_2, \ldots x_N) = 0. \end{aligned} \tag{16.1}$$

Ерөнхий тохиолдолд $\vec{F}(\vec{x}) = 0$ гэж илэрхийлдэг гэдгийг бид харсан. Цаашилбал

$$\vec{F}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_N(\vec{x}) \end{pmatrix} = 0. \tag{16.2}$$

Өмнө хэсэгт бид $N = 1$ үеийг сонирхосон. Одоо функц $f_i(\vec{x})$-ын Тейлорын цувааг сонирхоё.

$$f_i(\vec{x}^*) = f_i(\vec{x}^{(k)}) + \sum_{j=1}^{N} \frac{\partial f_i(\vec{x}^{(k)})}{\partial x_j} (\vec{x}^* - \vec{x}^{(k)}). \tag{16.3}$$

Иймээс ерөнхий систем тэгшитгэлийн хувьд

$$\vec{F}(\vec{x}^*) \approx \vec{F}(\vec{x}^{(k)}) + J(\vec{x}^{(k)})(\vec{x}^* - \vec{x}^{(k)}). \tag{16.4}$$

Энд $J(\vec{x}^{(k)})$-ийг нь Якобиан матриц гэдэг бөгөөд доорх байдлаар илэрхийлдэг,

Рис. 37. “Донровын Намдаг Төрийн хошой шагналт, УГЗ, Монголын орчин үеийн уран зохиолын гол төлөөлөгчдийн нэг, жүжгийн зохиолч, монголын театрын урлагийн анхдагчдын нэг.”

$$J(\vec{x}^{(k)}) = \begin{pmatrix} \frac{\partial f_1(\vec{x}^{(k)})}{\partial x_1} & \frac{\partial f_1(\vec{x}^{(k)})}{\partial x_2} & \cdots & \frac{\partial f_1(\vec{x}^{(k)})}{\partial x_N} \\ \frac{\partial f_2(\vec{x}^{(k)})}{\partial x_1} & \frac{\partial f_2(\vec{x}^{(k)})}{\partial x_2} & \cdots & \frac{\partial f_2(\vec{x}^{(k)})}{\partial x_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_N(\vec{x}^{(k)})}{\partial x_1} & \frac{\partial f_N(\vec{x}^{(k)})}{\partial x_2} & \cdots & \frac{\partial f_N(\vec{x}^{(k)})}{\partial x_N} \end{pmatrix}. \tag{16.5}$$

17 Ньютоны арга

Ньютоны аргын хувьд тэгшитгэл (16.4)-ийг ашиглан

$$\underbrace{\vec{F}(\vec{x}^*)}_{0} \approx \vec{F}(\vec{x}^{(k)}) + J(\vec{x}^{(k)})(\vec{x}^* - \vec{x}^{(k)}), \tag{17.1}$$

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - J^{-1}(\vec{x}^{(k)})\vec{F}(\vec{x}^{(k)}). \tag{17.2}$$

Бодит хэрэглээнд $J^{-1}(\vec{x}^{(k)})$-ийг тооцохын оронд

$$J(\vec{x}^{(k)})(\vec{x}^* - \vec{x}^{(k)}) = -\vec{F}(\vec{x}^{(k)}) \tag{17.3}$$

тэгшитгэлийг тооцон $\vec{y}^{(k)} = (\vec{x}^* - \vec{x}^{(k)})$-ийг олох бөгөөд $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \vec{y}^{(k)}$ гэж авч үздэг.

Algorithm 3 Ньютоны АРГА

ОРОЛТ:
$\vec{x}^{(0)}$ ▷ N хэмжээст анхны таамаг шийд вектор.
tol ▷ Хэрэв $\parallel \vec{x}^{(k+1)} - \vec{x}^{(k)} \parallel < tol$ бол алгоритм зогсоно.
N_{max} ▷ Хэрэв $N > N_{max}$ бол алгоритм зогсоно.
ГАРАЛТ:
ойролцоо шийд $\vec{x}^{(k)}$ ▷ тэгшитгэл $\vec{F}(\vec{x}^{(k)}) = 0$-ийг хангах.
$k = 0$
while $N \leq N_{max}$ do
    $\vec{F}(\vec{x}^{(k)})$ болон $J(\vec{x}^{(k)})$-ийг тооцох.
    $J(\vec{x}^{(k)})\vec{y}^{(k)} = -\vec{F}(\vec{x}^{(k)})$-ийг тооцох.
    $\vec{x}^{(k+1)} = \vec{x}^{(k)} + \vec{y}^{(k)}$
    if $\parallel \vec{y}^{(k)} \parallel < tol$ then
        break
    $k = k + 1$
Алгоритм Зогсох.

18 Шугаман Бус Оптимизац

Язгуур олох алгоритмыг ашиглан оптимизацын бодлогын шийдийг олж болох уу? Хамгийн бага утгыг нь олох шаардлагатай доорх шугаман бус бодлогыг авч үзэе.

$$\min_{\vec{x}} E(\vec{x}). \tag{18.1}$$

Мөн түүнчлэн $\max_{\vec{x}} E(\vec{x})$ шийдийг олох нь $\min_{\vec{x}}(-E(\vec{x}))$ шийд олохтой ижил юм. $E(\vec{x})$-ыг онцгой цэгд $\vec{x}^{(*)}$ шинжилхийн тулд

$$\nabla E(\vec{x}^*) = \begin{pmatrix} \frac{\partial E(\vec{x}^*)}{\partial x_1} \\ \frac{\partial E(\vec{x}^*)}{\partial x_N} \\ \vdots \\ \frac{\partial E(\vec{x}^*)}{\partial x_N} \end{pmatrix} = 0 \tag{18.2}$$

тэгшитгэлийг үнэлнэ. Онцгой цэг $\vec{x}^{(*)}$ нь $E(\vec{x})$-ын хамгийн бага цэг нь хэрэв доорх нөхцөл хангагддаг бол

$$\nabla^2 E(\vec{x}^*) = \begin{pmatrix} \frac{\partial^2 E(\vec{x}^*)}{\partial x_1 \partial x_1} & \frac{\partial^2 E(\vec{x}^*)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 E(\vec{x}^*)}{\partial x_1 \partial x_N} \\ \frac{\partial^2 E(\vec{x}^*)}{\partial x_2 \partial x_1} & \frac{\partial^2 E(\vec{x}^*)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 E(\vec{x}^*)}{\partial x_2 \partial x_N} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 E(\vec{x}^*)}{\partial x_N \partial x_1} & \frac{\partial^2 E(\vec{x}^*)}{\partial x_N \partial x_2} & \cdots & \frac{\partial^2 E(\vec{x}^*)}{\partial x_N \partial x_N} \end{pmatrix} > 0. \tag{18.3}$$

Тэгшитгэл (18.3) нь шугаман бус систем тэгшитгэлийг үүсгэдэг,

$$\vec{F}(\vec{x}) = \nabla E(\vec{x}^*) = 0. \tag{18.4}$$

Энэхүү тэгшитгэлийн шийдийг олохын тулд Ньютоны аргыг хэрэглэх боломжтой бөгөөд Якобиан матриц нь $J(\vec{x}^{(k)}) = \nabla^2 E(\vec{x}^{(k)})$ болно. Иймээс тэгшитгэл (17.3) нь

$$\begin{aligned} \nabla^2 E(\vec{x}^{(k)})\vec{y}^{(k)} &= -\nabla E(\vec{x}^{(k)}), \\ \vec{x}^{(k+1)} &= \vec{x}^{(k)} + \vec{y}^{(k)}. \end{aligned} \tag{18.5}$$

хэлбэрт орно. Ньютоны арга нь давталт бүрд 2-дугаар эрэмбийн уламжлалыг тооцох шаардлагатай болдог буюу иймээс бодит хэрэглээнд тохиромжгүй байдаг. Тиймээс

$$\begin{aligned} \vec{y}^{(k)} &= -\nabla E(\vec{x}^{(k)}), \\ \vec{x}^{(k+1)} &= \vec{x}^{(k)} + \eta \vec{y}^{(k)}, \end{aligned} \tag{18.6}$$

тэгшитгэлийг тооцох нь хялбар байдаг. Тэгшитгэл (18.6)-ийг Steepest Descent арга гэж нэрлэдэг.

Хүн-Машин

PART

IX

> Remark (Л. Түдэв, Оройгүй Сүм). Орчлон гаслангаас нөхцөгсдөд алагдана
> Зовлон их амгалан сахилд алагдана
> Амгалан билэг сэтгэлд алагдана

19 Интегралыг Тасралттай Функцын Аргаар Ойролцоолох

Энэ хэсэгт бид хэрхэн тодорхой интегралыг $[a, b]$ завсарт ойролцоолох талаар авч үзнэ,

$$I = \int_a^b f(x)dx; \tag{19.1}$$

учир нь зарим интегралын аналитик шийдийг олох нь зарим тохиолдолд боломжгүй эсвэл төвөгтэй, хэцүү түүнчлэн функц нь зөвхөн тодорхой цэгүүд дээр мэдэгдэж байдаг. Интеграл (19.1) -ийг ойролцоолохын тулд завсар $[a, b]$-ыг $a = x_0, x_1, x_2, \ldots, x_{N-1}, x_N = b$ хэсэгт хуваан, эдгээр цэгт доорх байдлаар ойролцоолон тооцох

$$I = \int_a^b f(x)dx \approx \sum_{i=0}^{N} w_i f(x_i), \tag{19.2}$$

цаашилбал бидний зорилго бол жин w_i-г тооцон олох юм. Интегралыг нийлбэрээр солидог нэгэн жишээг сонирхоё. Гадны тогтмол хүчний $F = const$ улмаас хийсэн ажил нь биеийг шилжүүлсэн замын өөрчилөлт $d = x_N - x_0$ болон үйлчилсэн хүчээр тодорхойлогддог, $A = F \cdot d$. Ажлын өөрчилөлт $\triangle A_i = F \cdot \triangle x_i$, болон нийт гүйцэтгэсэн ажил нь өөрчилөлтүүдийн нийлбэрээр $A = \sum_{i=0}^{N} \triangle A_i = \sum_{i=0}^{N} F \cdot \triangle x_i$. Хэрэв замын өөрчилөлтийг нь асар бага гэж авч үзвэл, эгэл замын хэсэг нь $dA_i = F \cdot dx_i$ гэж тодорхойлогддог буюу нийт гүйцэтгэсэн зам нь интегралиар $A = \int_{x_0}^{x_N} F dx$ илэрхийлэгддэг. Эндээс үзвэл хэрэв $N \to \infty$ ба $\triangle x_i \to 0$ бол,

$$\sum_{i=0}^{N} F \cdot \triangle x_i \to \int_{x_0}^{x_N} F dx. \tag{19.3}$$

Эсрэгээрээ N төгсгөлөг ба $\triangle x_i$ эгэл биш бол интеграл нь нийлбэр болон хувирдаг.

Рис. 38. “Гүн овогт Гун-Аажавын Аюурзана нь Монголын Улсын соёлын гавьяат зүтгэлтэн цолтой зохиолч юм. Тэрээр 1970 онд Баянхонгор аймагт төрсөн. 1988 – 1994 онд ОХУ-ын Москва хотноо М.Горькийн нэрэмжит Утга зохиолын дээд сургуульд суралцаж, дүүргэсэн. МЗЭ-ийн шагналт. 2002 оны шилдэг зохиолын "Алтан өд"цомын эзэн. Эхнэр Лувсандоржийн Өлзийтөгс нь мөн зохиолч. ”

20 Интегралыг Завсарт Хуваах

Интегралын утгын муж $[a, b]$-ийг N завсарт $[x_{i+1}, x_i]$ хуваан тооцно. Өөр үгээр илэрхийлбэл интегралын завсар $[a, b]$-ыг $[a, c]$ ба $[c, b]$ завсарт хуваавал,

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx. \tag{20.1}$$

Цаашилбал

$$\int_a^b f(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx \approx \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \tilde{F}(x)dx = \sum_{i=0}^{N-1} I_i. \tag{20.2}$$

энд функц $f(x)$-ыг тухайн завсарт ойролцоолон $\tilde{F}(x)$ гэж тэмдэглэв.

20.1 Тэгш өнцөгт, Трапец, Симпсоны дүрэм

Өөр өөр аргаар интегралыг ойролцоолохын тулд ашигладаг тоон аргууд нь функц $f(x)$-ыг ойролцоолох $\tilde{F}(x)$-ыг ялгаатай олон аргаар сонгодог.

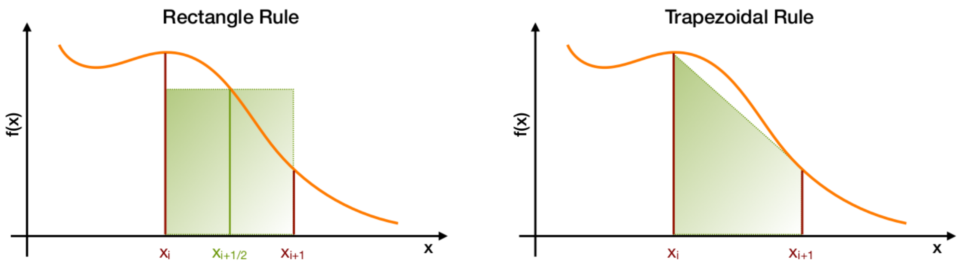


Рис. 39. Дурын матрицын Онцгой Утгын Задрал (SVD)

- Тэгш өнцөгтийн дүрэм: Хэрэв функц $f(x)$ нь $[x_i, x_{i+1}]$ мужид тогтмол бол тэгш өнцөгтийн дүрмийн дагуу (тэгши өнцөгтийн талбай нь өргөн, өндрийн үржвэртэй тэнцүү) энэ завсарт интеграл нь

$$I_{R_i} = f(x_{i+1/2})\triangle_i = f((x_i + x_{i+1})/2)\triangle_i.$$

- Трапецын дүрэм: Энэ тохиолдолд тэгш өнцөгт сонгохын оронд харин интегралыг трапецын талбай олохтой ижил авч үздэг. Ойролцоолж буй интеграл нь

$$I_{T_i} = \frac{f(x_i) + f(x_{i+1})}{2}\triangle_i.$$

- Симпсоны дүрэм: Тэгш өнцөгтийн болон трапецын дүрмийг нэмээд, нийт функцын тоо 3д хуваасан үр дүнг Симпсоны дүрэм гэдэг

$$I_{s_i} = I_{R_i} + I_{T_i} = \frac{f(x_i) + 4f((x_i + x_{i+1})/2) + f(x_{i+1})}{6}\triangle_i.$$

Функц $f(x)$-ыг x_i цэгүүдэд тооцохын тулд интегралын мужийг $[x_i, x_{i+2}]$, нийт цэг нь $N+1$ сондгой, мөн цэг хоорондох зай $\triangle_i$-г нь тогтмол гэж авч үзээ. Тэгвэл

$$I = \int_a^b f(x)dx = \sum_{i=0,i=even}^{N-1} \int_{x_i}^{x_{i+2}} f(x)dx,$$

$$\int_{x_i}^{x_{i+2}} f(x)dx \approx \frac{f(x_i) + 4f(x_{i+1}) + f(x_{i+2})}{3} \triangle_i.$$

болно.

Эцэст нь нэгтгэн илэрхийлбэл,

Тэгш өнцөгтийн дүрэм:

$$I = \triangle \sum_{i=0}^{N-1} f(x_i), \tag{20.3}$$

Трапецын Дүрэм:

$$I \approx \frac{\triangle}{2} \left(f_0 + 2 \sum_{i=1}^{N-1} f_i + f_N \right), \tag{20.4}$$

Симпсоны Дүрэм:

$$I \approx \frac{\triangle}{3} \left(f_0 + 4 \sum_{i=1,i=odd}^{N-1} f_i + 2 \sum_{i=2,i=even}^{N-2} f_i + f_N \right). \tag{20.5}$$

20.2 Newton-Cotes formula

Интегралыг ойролцоолох ерөнхий арга бол Ньютон-Котесын тэгшитгэл юм. Тэгшитгэл (20.3)-д өгөгдсөн ойролцоо функц $\tilde{F}_i(x)$-ыг байгуулахын тулд $[x_{i+1}, x_i]$ завсарыг $M+1$ тэнцүү завсарт $(x_k = x_i + k \cdot h, \quad k = 0, \dots M)$ хуваах бөгөөд Лагранжийн интерполяцын аргыг ашиглана. $x_k, f(x_k)$ цэгүүдийг дайрч гарсан Лагранж интерполяц нь

$$\begin{aligned} \tilde{F}_i(x) &= \sum_{k=0}^{M} f(x_k) l_k^M(x), \\ l_k^M(x) &= \frac{(x-x_0)(x-x_1)\dots(x_k-x_M)}{(x_k-x_0)(x_k-x_1)\dots(x_k-x_M)}. \end{aligned} \tag{20.6}$$

Иймээс интеграл (20.2) нь доорх хэлбэрт орно

Ньютон-Котесын тэгшитгэл:

$$
\begin{aligned}
I_i = \int_{x_i}^{x_{i+1}} f(x) \approx \int_{x_i}^{x_{i+1}} \tilde{F}_i(x) &= \triangle \sum_{k=0}^{M} f(x_k) \underbrace{\frac{1}{\triangle} \int_{x_i}^{x_{i+1}} l_k^M(x)}_{C_k^M(x)}, \\
&= \triangle \sum_{k=0}^{M} f(x_k) C_k^M(x).
\end{aligned}
\tag{20.7}
$$

Хэрэв $M = 1$ бол трапецын дүрмийг, харин $M = 2$ бол Симпсоны дүрмийг гаргаж авах боломжтой гэдгийг хялбархан харуулж болно.

PART

X

Хүн-Машин

Remark (Л. Түдэв, Оройгүй Сум). Хуримд хувцасгүй хүүхэн үнэгүй
Хуралд номгүй банди үнэгүй
Худалдаанд шүдгүй тэмээ үнэгүй

21 Ромберг Интеграл дүрэм

Өмнөх хэсэгт дурдсан интегралыг ойролцоолох Ньютон-Котесын тоон арга нь олон онцгой цэгтэй функцын хувьд тохиромжгүй байдаг. Харин нэгэн тохиромжтой аргыг Ромбергийн интегралын арга гэдэг буюу энэхүү арга нь Ричардсоны экстреполяцыд суурилдаг.

Рис. 40. “1972 онд Дархан хотод төрсөн. 1996 онд Монгол мэдлэгийн их сургуулийг судлаач мэргэжилээр төгссөн. 2002 онд "Эрх чөлөөтэй байхын урлаг буюу шинэ ном"номоороо Алтан өд шагнал, 2004 онд мөн энэ номоороо МЗЭ -ийн шагнал авчээ. ”

21.1 Ричардсоны экстреполяц

Компьютерд бүх зүйлийг тасралттай тоогоор илэрхийлдэг (жнь: зураг, дүрст бичлэг). Коипьютер ашиглан тооцох функцын нарийвчилал эсвэл компьютерд хадгалсан тухайн сонирхож буй зүйлийн нүдэнд харагдах чанар нь тасралттай тооноос h хамаардаг бөгөөд жишээ болгон доорх зургийг сонирхоё,

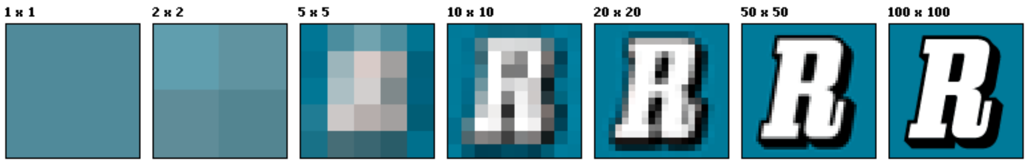


Рис. 41. Зургийн өргөн ба өндөр нь $1cm$: Эхний зурагт пикселийн өргөн нь $h = 1cm$, 2 дах зурганд $h = 0.5cm$, гэх мэтчилэн нэг дахин багасгавал зургийн нарийвчилал болон нүдэнд харагдах чанар сайжирч байна (source: Wikipedia).

Үүнээс үзэхэд тасралттай тоо h нь зургийн чанарыг сайжруулж байна. Хэрэв бид дээрх зургийг F гэж тэмдэглэвэл, F-ийн нарийвчилал нь пикселийн тоо буюу h-ээс хамаарч байгаа учир математикийн хувьд доорх байдлаар илэрхийлж болно,

$$F \approx F(h). \tag{21.1}$$

Өөр үгээр илэрхийлбэл комьпютерд оруулж буй тухайн зүйлийн нарийвчилал нь сонгож авсан h-ээс хамаардаг. Сонгож авсан h нь маш бага үед $h << 1$ бид (21.1)-ыг тэгийн орчим Тейлорын цуваанд задлах боломжтой

$$F(h) = F(0) + F^{'}(0)h + F^{''}(0)\frac{h^2}{2}\ldots, \tag{21.2}$$

өмнөх хэсгүүдэд дурдсанчилан $F = F(0)$ нь жинхэнэ утга харин нэмэлт хэсгүүд h нь алдааны утгууд байдаг. Тооцон бодох инженерийн салбарт функцыг тооцох хурд ба нарийвчилал чухал байдаг учир тэгшитгэл (21.2)-ийг тооцохын тулд $h = h/2$ гэж авч үзээ, тэгвэл

$$F(h/2) = F(0) + F^{'}(0)\frac{h}{2} + F^{''}(0)\frac{h^2}{8}\dots. \tag{21.3}$$

Ричардсоны гол санаа нь тэгтшигэл (21.2) ба (21.3)-ийг нэгтгэх байсан бөгөөд

$$F_1(h) = 2F(h/2) - F(h) = F(0) + c_2h^2 + c_3h^3 + \dots, \tag{21.4}$$

энд c_2, c_3 нь функцын уламжлалтай хэсгүүд болно. Энд алдааны утга нь багасч h^2 байна гэдгийг харж болно. Харин (21.2) болон (21.3)-т h байсан. Жишээ нь хэрэв бидэнд (21.2) болон (21.3)-ын хувьд $h = 0.1$ байсан бол (21.4) хувьд $h^2 = 0.01$ болж багасч байна. Тэгшитгэл (21.4)-д хэрэглэсэн арыг цааш давтван илэрхийлбэл

$$F_2(h) = \frac{1}{3}(4F_1(h) - F_1(h)) = F(0) + \dots. \tag{21.5}$$

Эцэст нь ерөнхий тэгшитгэлийг бичвэл

$$F_n(h) = \frac{1}{2^n - 1}(2^nF_{n-1}(h/2) - F_{n-1}(h)) = F(0) + \dots. \tag{21.6}$$

21.2 Жишээ бодлого

Функцын уламжлалыг доорх байдлаар ойролцоолж болдог

$$f^{'}(x) \approx \frac{f(x+h) - f(x)}{h} = F_0(h) \tag{21.7}$$

бөгөөд уламжлалын утгыг $F_0(h)$ гэж үзээ. Тэгвэл ойролцоолж буй уламжлалын утгын нарийвчлалыг сайжруулахын тулд тэгшитгэл (21.6)-г ашиглах буюу дурын функц

$$f(x) = x + e^x$$

-ыг сонгон авч $x = 0$ цэг дэх уламжлалыг тэгшитгэл 6-ын дагуу $F_2(h)$-ийг олъё. Тэгшитгэлийн бодит шийд нь $f^{'}(0) = 2.00$, $h = 0.4$ гэж өгөгдсөн гэж үзээ.

$$
\begin{aligned}
F_1(h) &= \frac{1}{2}(2F_0(h/2) - F_0(h)) \\
&= \left[2 \cdot \frac{((x+0.2)+e^{x+0.2}) - (x+e^x)}{0.2} - \frac{((x+0.4)+e^{x+0.4}) - (x+e^x)}{0.4}\right]; \\
F_2(h) &= \frac{1}{3}(4F_1(h/2) - F_1(h)) \\
&= \frac{1}{3} \cdot 4 \cdot \left[2 \cdot \frac{((x+0.1)+e^{x+0.1}) - (x+e^x)}{0.1} - \frac{((x+0.2)+e^{x+0.2}) - (x+e^x)}{0.2}\right] \\
&- \frac{1}{3}\left[2 \cdot \frac{((x+0.2)+e^{x+0.2}) - (x+e^x)}{0.2} - \frac{((x+0.4)+e^{x+0.4}) - (x+e^x)}{0.4}\right] \\
&= 2.00039.
\end{aligned}
\tag{21.8}
$$

Тийм бол энэхүү олсон шийдийг бодит шийдтэй нь харцуулж алдааг нь олбол $e = |2.00039 - 2.00| = 0.00039$ байна. Хэрэв Ричардсоны аргыг үл ашиглан зөвхөн тэгшитгэл (21.7)-ийн дагуу шийдийг нь олсон бол шийд нь $F_0(h) = 2.22956$, алдаа нь $e = |2.22956 - 2.00| = 0.22956$ байна. Үүнээс үзвэл Ричардсоны арга нь ойролцоо шийдийн нарийвчлалыг ихэсгэж байна.

22 Ромберг Интегралчилал

Ромбергийн арга нь Ричардсоны аргад суурилсан бөгөөд зорилго нь интегралын ойролцоо шийдийн нарийвчлалыг сайжруулах юм. Трапецын дүрмийг жишээ болгон авч үзээ. Ричардсоны аргатай адил сонгон авч буй завсрын өргөн h-ийг нэг дахин n удаа багасгах замаар интегралыг ойролцоолно, (Рис. 2). Өөр үгээр илэрхийлбэл интегралыг $I_0^1, I_0^2, \ldots, I_0^n$ гэх мэтчилэн өөр өөр завсарт тооцно.

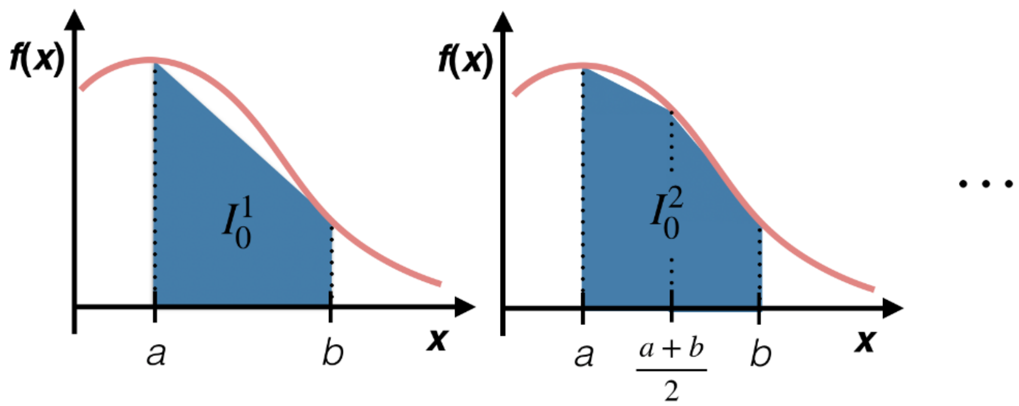


Рис. 42. Trapezoidal Rule For Romberg Integration

Трапецын интеграл дүрмийг доорх байдлаар илэрхийлж болно,

Algorithm 4 Ромберг Интегралчилал АРГА

ОРОЛТ:
Функц $f(x)$
Завсар $[a, b]$
Давталтын тоо K
ГАРАЛТ:
$I_K^1 = integral[K, 0]$ ▷ $\int_a^b f(x)dx$ интегралын ойролцоолол.
$maxNumInt \leftarrow 2^k$
$hmin \leftarrow (b-a)/maxNumInt$
for $i \leftarrow 0, \dots, maxNumInt$ do
    $fvalues[i] \leftarrow f(a + i \cdot hmin)$
for $r \leftarrow 0, \dots, K$ do
    $numInt \leftarrow 2^r$
    $step \leftarrow 2^{K-r}$
    $result \leftarrow 0$
    for $i \leftarrow step, 2 \cdot step, 3 \cdot step, \dots, maxNumInt - step$ do
        $result \leftarrow result + fvalues[i]$
    $integral[0, r] \leftarrow 0.5\frac{b-a}{numInt}(fvalues[0] + fvalues[maxNumInt] + 2 \cdot result)$
for $l \leftarrow 1, \dots, K$ do
    for $r \leftarrow 0, \dots, K - l$ do
        $integral[l, r] \leftarrow \frac{4^l \cdot integral[l-1, r+1] - integral[l-1, r]}{4^l - 1}$
Алгоритм Зогсох.

$$I = \int_a^b f(x)dx = \underbrace{\frac{h}{2}\left[f(a) + f(b) + 2\sum_{i=1}^{n-1} f_i\right]}_{I_0^n} + c_1h^2 + c_2h^4 + \dots \tag{22.1}$$

энд $h = (b-a)/n$, $f_i = f(a + ih)$.

$$I_0^n = I - c_1h^2 - c_2h^4 - c_3h^6 - \dots \tag{22.2}$$

Сонгож авсан завсарыг нэг дахин багасгавал $h = h/2$.

$$I_0^{2n} = I - c_1\frac{h^2}{4} - c_2\frac{h^4}{16} - c_3\frac{h^6}{64} - \dots \tag{22.3}$$

Ричардсоны экстреполяцын арга (21.6)-г ашиглан (22.1)-ийг илэрхийлбэл

$$I_1^n = \frac{4I_0^{2n} - I_0^n}{3} = I + c_2\frac{h^4}{16} + c_3\frac{5h^6}{16} + \dots \tag{22.4}$$

Дахин тэгшитгэл (22.2)-ийн алхмыг 2 дахин багасгабал $h = h/4$,

$$I_0^{4n} = I - c_1\frac{h^2}{16} - c_2\frac{h^4}{256} - c_3\frac{h^6}{4096} - \dots \tag{22.5}$$

мөн Ричардсоны экстреполяцын арга (21.6)-г ашиглабал,

$$I_1^{2n} = \frac{4I_0^{4n} - I_0^{2n}}{3} = I + c_2\frac{h^4}{64} + c_3\frac{5h^6}{1024} + \dots \tag{22.6}$$

гэх мэтчилэн энэ үйлдлийг давтвал интегралын ойролцоо шийд нь

$$I_k^{2n} = \frac{4^k I_{k-1}^{2n} - I_{k-1}^{n}}{4^k - 1}. \tag{22.7}$$

22.1 Жишээ бодлого

Энд интеграл $err(t) = 2/\sqrt{\pi}\int_0^{0.5} e^{-x^2}dx$-ын утгыг тооцоё. Энэхүү интегралын жинхэнэ шийд нь $err = 0.5204998778130467$ бөгөөд бидний тооцсон шийд нь $err = 0.5204998778129182$ байна. Эндээс үзвэл Ромбергийн интегралчилах арга нь шийдийн нарийвчилалыг ихэсгэж байна гэдэг нь харагдана.

In [4]:
```python
from scipy import integrate
from scipy.special import erf
import numpy as np
```

Бид функц $err = 2/\sqrt{(\pi)}\int_0^{0.5} e^{-t^2}dt$-ын утгыг сонирхоё. Тэгвэл доорх байдлаар үр дүн нь олдоно:

In [6]:
```python
err = lambda t: 2/np.sqrt(np.pi) * np.exp(-t**2)
result = integrate.romberg(err, 0, 0.5, show=True)
```

```
Romberg integration of <function vectorize1.<locals>.vfunc at 0x7f096aeb18c8> from [0, 0.5]

 Steps  StepSize   Results
     1  0.500000  0.501790
     2  0.250000  0.515899  0.520602
     4  0.125000  0.519354  0.520506  0.520500
     8  0.062500  0.520214  0.520500  0.520500  0.520500
    16  0.031250  0.520428  0.520500  0.520500  0.520500  0.520500

The final result is 0.5204998778129182 after 17 function evaluations.
```

Рис. 43. Romberg Integration: python package.

PART

XI

Хүн-Машин

> Remark (Франц Кафка). Битгий гудай, Битгий хүндээр хүлээж ав
> Битгий рационал байх гэж хичээ
> Битгий өөрөө, өөрийгөө бусдад тааруулж өөрчил
> Харин өөрийн хамгийн хүчтэй хүсэл зорилгоо дага

Рис. 44. “Франц Кафка (Герман: Franz Kafka; 1883 оны 7 сарын 3 – 1924 оны 6 сарын 3) нь XX зууны Герман хэлтний уран зохиолын томоохон төлөөлөгч байлаа. Чехийн Праг хотноо худалдаачин еврейн гэр бүлд төрсөн. Эх нь эрдэмт раббайн гаралтай, эцэг нь мал нядлагчийн хүү байжээ.”

23 Adaptive Integration

Ричардсоны экстреполяцын хувьд функц $F(0)-$ын ойролцоо шийдийг $h = h/2$ үед доорх байдлаар тодорхойлох боломжтой гэдгийг авч үзсэн.

$$F(h/2) = F(0) + F^{'}(0)\frac{h}{2} + F^{''}(0)\frac{h^2}{8}. \tag{23.1}$$

Тэгвэл энэ үед тэгшитгэлийн алдаа нь

$$\epsilon(h/2) = F(h/2) - F(0) = F^{'}(0)\frac{h}{2} + F^{''}(0)\frac{h^2}{8}, \tag{23.2}$$

гэж тодорхойлогдоно. Иймээс алхамын хуваалт бүрд алдааг нь дээрх маягаар тооцох боломжтой. Өмнөх хэсэгт функцын интегралыг тооцохын тулд алхамын өргөн h нь бүх мужид ижил байсан. Гэвч зарим интегралын хувьд муж бүрийг өөр, өөр алхамд хуваах шаардлагатай болдог. Жишээ нь доорх функцын талбайг трапецын дүрмийг ашиглан олоё гэж үзэе.

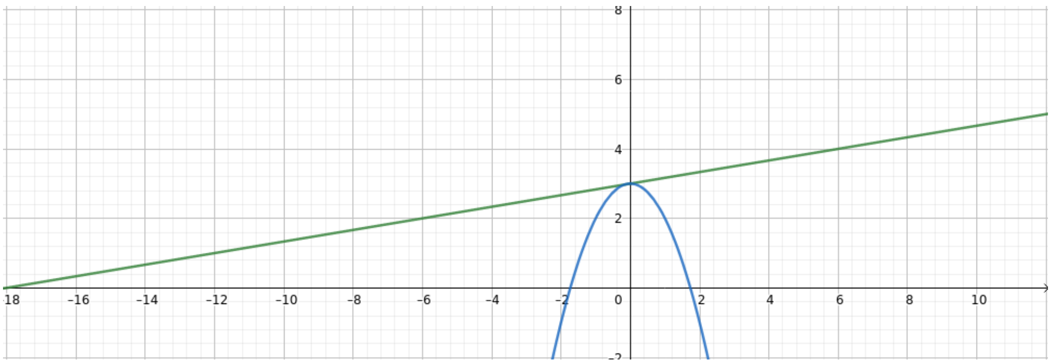


Рис. 45. F функц гэж үзэе.

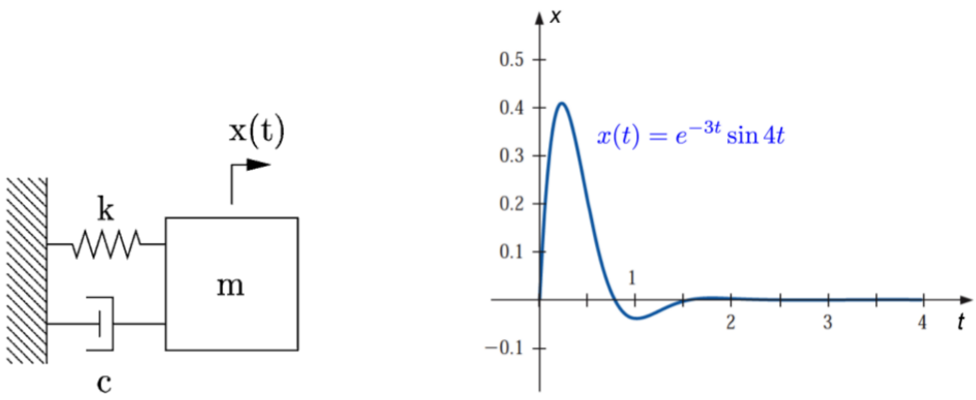
Энэхүү жишээг авч үзвэл, $[-2, +2]$ мужийг л маш бага олон алхамд h хуваах шаардлагатай, харин бусад мужийг олон алхамд h хуваах шаардлагагүй гэдэг нь харагдаж байна. Энэхүү аргыг Adaptive Integration гэж нэрлэдэг.

23.1 Жишээ бодлого

Adaptive Integration аргыг унтрах хэлбэлзлийн шийдийг олоход ашиглая. Унтрах хэлбэлзэлийн хөдөлгөөний тэгшитгэл нь $m\ddot{x}(t) + c\dot{x}(t) + kx(t) = 0$, $m = 1, c = 6, k = 25$, $x(0) = 0, \dot{x}(0) = 4$ бөгөөд шийд нь $x(t) = e^{-3t}\sin(4t)$ гэж тодорхойлогдох бол хугацаа нь $t = 0$-оос $t = 4$ үед нийт туулсан замыг нь аналитикаар олбол

$$\int_0^4 e^{-3t}\sin(4t)dt \approx 0.160001153722807, \tag{23.3}$$

болно. Одоо тэгвэл тэгшитгэл (23.3)-ын шийдийг Adaptive Integration аргыг



(a) (b)

Рис. 46. Spring-Damper Oscillation (source ocw.mit.edu, ETH).

ашиглан олж дээр олсон бодит үр дүнтэй нь харьцуулая. Алгоритм 1-ийн хувьд

Algorithm 5 AdaptiveTrapezoidal

Функц: AdaptiveTrapezoidal(a,b) ▷ интегралын хил $[a, b]$.
Алдааны нарийвчилал $tol = 10^{-6}$
Трапецын интегралчилах дүрмийг $[a, b]$ завсарт хэрэглэнэ.
Интегралын завсрыг 2 хуваана: $[a, m]$ мөн $[m, b]$, энд $m = (a + b)/2$.
Трапецын интегралчилах дүрмийг завсар $[a, m]$ мөн $[m, b]$-т хэрэглэнэ.
Ричардсоны экстреполяцын аргыг (23.2) ашиглан $[a, b]$ алдааг нь тооцно.
if хүссэн нарийвчилал tol-д хүрээгүй бол **then**
    return AdaptiveTrapezoidal(a,m)+AdaptiveTrapezoidal(m,b)
return AdaptiveTrapezoidal(a,m)+AdaptiveTrapezoidal(m,b)
гаралтын утга: $(0.108659, 0.160001)$

гаралтанд өгсөн эхний утга нь трапецын интегралын дурмийг нэг удаа ашиглахад, харин сүүлийнх нь Adaptive Trapezoidal дүрмийг ашиглаж олж авсан утгууд болно. Үүнээс үзвэл Adaptive Trapezoidal аргыг ашиглан олсон шийд нь бодит шийд (23.3)-тай ойролцоо олдсон байна гэдэг нь харагдсаж байна.

24 Gauss Quadrature

Бидэнд дурын функцын ойролцоо интгеграл нь доорх байдлаар өгөгдсөн гэж үзээ,

$$\int_a^b f(x)dx \approx c_1 f(a) + c_2 f(b). \tag{24.1}$$

Хэрэв $f(x) = a_0 + a_1 x$ гэж өгөгдсөн бол, тэгшитгэл (23.3)-ын зүүн гар тал нь

$$\int_a^b f(x)dx = \int_a^b (a_0 + a_1 x)dx = a_0(b-a) + a_1\left(\frac{b^2 - a^2}{2}\right). \tag{24.2}$$

Цаашилбал (24.1)-ийн баруун гар талыг (23.3)-тай харьцуулбал,

$$a_0(b-a) + a_1\left(\frac{b^2-a^2}{2}\right) = c_1(a_0 + a_1 a) + c_2(a_0 + a_1 b). \tag{24.3}$$

(24.2)-ын баруун гар талыг хялбарчилбал $a_0(c_1+c_2)+a_1(c_1 a+c_2 b)$ болох бөгөөд эндээс

$$c_1 = c_2 = \frac{b-a}{2}, \tag{24.4}$$

гэж олдох буюу (23.3)-т орлуулбал интегралыг ойролцоолох трапецын дүрмийг гарган авч болж байна.

24.1 Хоёр-Цэгийн Gauss quadrature

Цаашилбал интеграл (23.3)-ын баруун гар тал нь үл мэдэгдэх 2-цэгээр тодорхойллогддог гэж үзвэл,

$$\int_a^b f(x)dx \approx c_1 f(x_0) + c_2 f(x_1). \tag{24.5}$$

Кубик функц $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$-ыг ашиглан,

$$\begin{gathered}\int_a^b (a_0 + a_1 x + a_2 x^2 + a_3 x^3)dx = \\ a_0(b-a) + a_1\left(\frac{b^2-a^2}{2}\right) + a_2\left(\frac{b^3-a^3}{3}\right) + a_3\left(\frac{b^4-a^4}{4}\right).\end{gathered} \tag{24.6}$$

өмнө дурдсан (24.1)-(24.3)-тай ижил алхмаар (24.4)-ийн үл мэдэгдэх коэффициентүүдийг тодорхойлоё. Тэгвэл

$$\begin{aligned} c_1 &= \frac{b-a}{2}, \\ c_2 &= \frac{b-a}{2}, \\ x_0 &= \left(\frac{b-a}{2}\right)\left(-\frac{1}{\sqrt{3}}\right) + \frac{b+a}{2}, \\ x_1 &= \left(\frac{b-a}{2}\right)\left(\frac{1}{\sqrt{3}}\right) + \frac{b+a}{2}; \end{aligned} \tag{24.7}$$

болох бөгөөд энэхүү аргыг Хоёр-Цэгийн Gauss quadrature арга гэж хэлдэг.

24.2 n-цэгийн Gauss quadrature

24.2.1 Хермитийн полином

Өмнөх хэсгүүдэд өгөгдсөн өгөгдөлд (x_i, y_i) харгалзах зэрэгт функцын интерполяц нь $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ гэж тодорхойлогдсон буюу $p(x_i) = y_i$ хамаарлыг олох нь гол зорилго байдаг. Харин Хермитийн полиномиал интерполяц нь өгөгдөлийн уламжлалын мэдээллийг мөн ашигладаг бөгөөд доорх байдлаар тодорхойлогддог,

$$f(x_i) = y_i, f^{'}(x_i) = y_i^{'}, \tag{24.8}$$

энд $i = 1, 2, 3, \ldots, n$. Иймээс бидний зорилго бол (24.7)-ыг ойролцоолох $2n - 1$ эрэмбийн полиномиал функцыг $f(x)$ олох юм,

$$f(x) = \sum_{k=1}^{n} U_k(x) y_k + \sum_{k=1}^{n} V_k(x) y_k^{'}. \tag{24.9}$$

$U_k(x), V_k(x)$ нь (24.7)-ийг хангахын тулд доорх шинж чанартай,

$$\begin{aligned} &U_k(x_j) = \delta_{kj}, \quad U_k^{'}(x_j) = 0 \\ &V_k(x_j) = 0, \quad V_k^{'}(x_j) = \delta_{kj}. \end{aligned} \tag{24.10}$$

Иймээс бид Лагранжийн полиномыг ашиглан тэгшитгэл (24.9)-ыг илэрхийлэх боломжтой.

$$\begin{aligned} L_k(x) &= \frac{(x - x_1)(x - x_2)\ldots(x - x_n)}{(x_k - x_1)(x_k - x_2)\ldots(x_k - x_n)}, \\ U_k(x) &= \left(1 - 2L_k^{'}(x_k)(x - x_k)\right) L_k^2(x), \\ V_k(x) &= (x - x_k) L_k^2(x). \end{aligned} \tag{24.11}$$

Тэгшитгэл (24.8)-ийг Хермитийн полином гэж нэрлэдэг.

24.2.2 n-цэгийн Gauss quadrature

n-цэгийн Gauss quadrature дүрмийг $[-1, +1]$ мужид тооцох нь тохиромжтой байдаг, учир нь энэ мужид функцыг ойрооцоолж буй полином (24.8) нь нарийвчилал өндөртэй байдаг. Иймээс бид интегралын мужийг $[a, b]$,

$$\epsilon = \frac{2x - (a + b)}{b - a}, \tag{24.12}$$

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2-1)$
3	$\frac{1}{2}(5x^3-3x)$
4	$\frac{1}{8}(35x^4-30x^2+3)$
5	$\frac{1}{8}(63x^5-70x^3+15x)$
6	$\frac{1}{16}(231x^6-315x^4+105x^2-5)$
7	$\frac{1}{16}(429x^7-693x^5+315x^3-35x)$
8	$\frac{1}{128}(6435x^8-12012x^6+6930x^4-1260x^2+35)$
9	$\frac{1}{128}(12155x^9-25740x^7+18018x^5-4620x^3+315x)$
10	$\frac{1}{256}(46189x^{10}-109395x^8+90090x^6-30030x^4+3465x^2-63)$

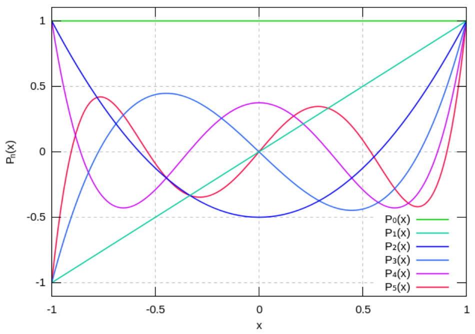


Рис. 47. Лежендр полином.

ашиглан $[-1,+1]$ мужид шилжүүлэх хэрэгтэй.

$$
\begin{aligned}
\int_{-1}^{+1} f(x)dx &= \sum_{k=1}^{n} y_k \int_{-1}^{+1} U_k(x)dx + \sum_{k=1}^{n} y_k^{'} \int_{-1}^{+1} V_k(x)dx,\\
&= \sum_{k=1}^{n} u_k f(x_k) + \sum_{k=1}^{n} v_k f^{'}(x_k), \qquad (24.13)\\
u_k &= \int_{-1}^{+1} U_k(x)dx, \quad v_k = \int_{-1}^{+1} V_k(x)dx.
\end{aligned}
$$

Интегралыг $\int_{-1}^{+1} f(x)dx = \sum_{i=1}^{n} w_i f(x_i)$ хэлбэрт бичихийн тулд $v_k = 0$ байх ёстой гэдгийг (24.12)-аас хялбархан ажиглаж болно. Цаашилбал бид бяцхан математик үйлдлийг гүйцэтгэвэл доорх үр дүнд хүрэх бөгөөд

$$
\begin{aligned}
v_k &= \int_{-1}^{+1} (x-x_k) L_k^2(x)dx,\\
L_k(x) &= \frac{\underbrace{(x-x_1)(x-x_2)\dots(x-x_n)}_{F(x)}}{\underbrace{(x_k-x_1)(x_k-x_2)\dots(x_k-x_n)}_{C_k}}, \qquad (24.14)\\
v_k &= C_k \int_{-1}^{+1} F(x)L_k(x)dx,\\
0 &= \int_{-1}^{+1} F(x)L_k(x)dx, \quad C_k \neq 0.
\end{aligned}
$$

хангах x_k-үүдийг тодорхойлох хэрэгтэй болно гэсэн үг юм. $F(x) = P_n(x) = (x-x_1)(x-x_2)\dots(x-x_n)$-ийг Лежендрийн полином гэж нэрлэдэг (Зураг 2). Эцэст нь n-цэгийн Gauss Quadrature нь доорх байдлаар тодорхойлогдоно,

$$
\int_{-1}^{+1} f(x)dx = \sum_{k=1}^{n} u_k f(x_k), \quad u_k = \frac{2}{(1-x_k^2)(P_n^{'}(x_k))^2}. \qquad (24.15)
$$

PART

XII

Хүн-Машин

> Remark (Мүраками Харүки). Хэрэв та бусад хүмүүсийн уншсан номыг л уншдаг бол
> та тэдгээр бусадтай ижил үзэл бодол, сэтгэлгээтэй л байх болно

25 Curse of Dimensional

Бидэнд $d \in \mathbb{N}$ ширхэг хувьсагчаас хамаарч буй $f : \mathbb{R}^d \to \mathbb{R}, f : \vec{x} = \{x^{(1)}, x^{(2)}, \dots x^{(d)}\} \to f(\vec{x})$ функц өгөгдсөн гэж үзэе. Сонирхож буй интегралын утгын муж нь $\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_d$, $\Omega_i = [a_i, b_i]$ гэж өгөгдсөн бол,

$$I = \int_\Omega f(\vec{x}) d\vec{x} = \int_{\Omega_1} \int_{\Omega_2} \cdots \int_{\Omega_d} f(\vec{x}) dx^{(d)} \dots dx^{(2)} dx^{(1)}. \tag{25.1}$$

Дээрх интегралыг тооцохын тулд n-цэгийн Gauss quadrature дүрмийг интеграл бүрд $i = 1, \dots, d$ ашиглах боломжтой. Нэг хэмжээст Gauss quadrature дүрмийг $x_1, x_2, \dots, x_n$ цэгүүд болон $w_1, w_2, \dots, w_n$ жинтэй тохиолдолд ашиглавал, d хэмжээст интеграл нь

$$I \approx \sum_{i_1=1 \dots i_n=1}^{n} \tilde{w}_{i_1,\dots i_d} f(x_{i_1}^{(1)}, \dots x_{i_d}^{(d)}), \quad \tilde{w}_{i_1,\dots i_d} = \prod_{r=1}^{d} w_{i_r}. \tag{25.2}$$

Энд нийт үнэлэх функцын тоо нь $N = n^d$ болж байна. Жишээ болгон $d = 2, n = 3$ үед доорх функцыг сонирхож үзвэл,

$$\begin{aligned} I &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x,y) dx dy = \sum_{j=1}^{3} \sum_{i=1}^{3} w_i w_j f(x_i, y_j) \\ &= \sum_{j=1}^{3} (w_1 w_j f(x_1, y_j) + w_2 w_j f(x_2, y_j) + w_3 w_j f(x_3, y_j)) \\ &= (w_1 w_1 f(x_1, y_1) + w_2 w_1 f(x_2, y_1) + w_3 w_1 f(x_3, y_1)) \\ &+ (w_1 w_2 f(x_1, y_2) + w_2 w_2 f(x_2, y_2) + w_3 w_2 f(x_3, y_2)) \\ &+ (w_1 w_3 f(x_1, y_3) + w_2 w_3 f(x_2, y_3) + w_3 w_3 f(x_3, y_3)) \,. \end{aligned} \tag{25.3}$$

буюу $N = n^d = 3^2 = 9$ гэдэг нь харагдаж байна. Одоо тэгвэл curse of dimensional гэж ямар утгатай вэ? гэдгийг сонирхоё. Өмнөх хэсэгт үзсэн нэг хэмжээст Симпсоны дүрмийн хувьд интегралын нарийвчилал нь алхмын өргөнийг 4 зэрэгт дэвшүүлсэнтэй тэнцүү h^4 гэдгийг үзсэн бөгөөд $(N = n, d = 1)$ үед

$$I - I_s = O(h^4), \quad h = \frac{b-a}{n} = O(n^{-1}) = O(N^{-1}), \tag{25.4}$$

Рис. 48. “Мүраками Харүки (1949 оны 1 сарын 12-нд төрсөн) нь Японы зохиолч, орчуулагч юм. Түүний утга зохиолын болон баримтат зохиолын бүтээлүүд нь Франц Кафкагийн шагнал, Франц О Коннорын богино өгүүллэгийн шагнал, Иерусалимын шагнал зэрэг олон улсын утга зохиоын шагналуудыг хүртсэн юм.”

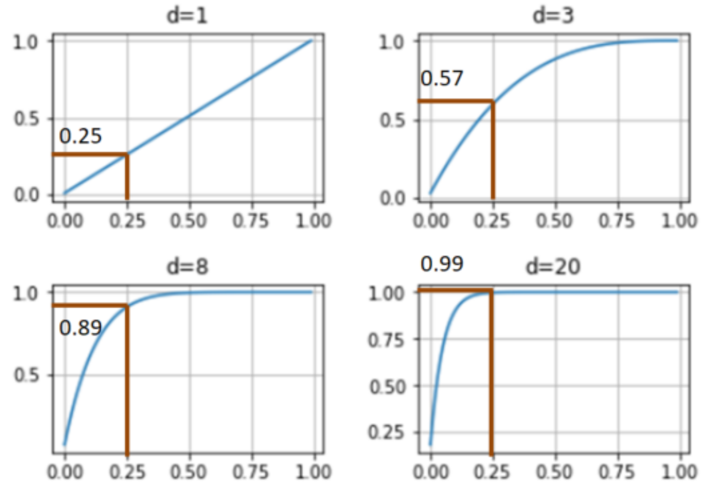


Рис. 49. Энэхүү графикийн $\epsilon = 0.25$ үед харгалзах эзэлхүүний хувийг өөр өөр хэмжээст огторгуйд тооцон тэмдэглэсэн болно.

дээрх байдлаар илэрхийлж болно. Цаашилбал d хэмжээст үед

$$h = O(n^{-1}) = O(N^{-1/d}). \tag{25.5}$$

Эндээс үзвэл хэмжээс d-г ихэсгэх тусам, нарийвчилал h нь буурч байна. Үүнийг curse of dimensional гэж нэрлэгдэг буюу d хэмжээст интегралыг тооцох өөр арга хэрэгтэй гэдэг нь харагдаж байна. Энэхүү аргыг Монте Карло гэдэг.

Мөн түүнчилэн curse of dimensional хэмээх ойлголт машин сургалтын өгөгдөлд ажиллах үед чухал хэрэгтэй байдаг. Учир нь n тооны өгөгдөлийн $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)$ хэмжээс ихсэх тусам машин сургалтын алгоритмыг тооцолоход хүдрэлтэй болдог. Жишээ болгож $d = 3$ хэмжээст огторгуйд орших бөмбөрцөгийн эзэлхүүнийг авч үзэе. Өөр үгээр хэлбэл өгөгдөл нь $d = 3$ хэмжээст огторгуйд орших цэгүүд юм; $\vec{x}_i = (x_i, y_i, z_i)^T$. Бидний мэдхээр бөмбөрцөгийн эзэлхүүн нь $d = 3$ үед $V_3 = (4\pi/3)r^3 = c \cdot r^3$, төдийгүй ерөнхий d тохиолдолд бөмбөрцөгийн эзэлхүүн $V_d = k \cdot r^d$ гэж тодорхойлогддог. Жишээ болгон r_1 мужид орших $r - \epsilon \le r_1 \le r$ бөмбөрцөгийн эзэлхүүний эзлэх хувийг тодорхойлоё. Тэгвэл $r = 1$, $V = k \cdot r^d$, $V_e = k \cdot (r - e)^d$ гэж авч үзэе;

$$\frac{V - V_e}{V} = 1 - (1 - \epsilon)^d. \tag{25.6}$$

Одоо $d = 1, 3, 8, 20$ үеийн графикийг $\epsilon = 0.0, 0.01, 0.02, \ldots, 1$ үед байгуулая (Зур. 1). Дээрх зургаас үзвэл хэмжээс d ихсэх тусам нийт эзэлхүүн нь бөмбөрцөгийн гадаргын ойролцоо төвлөрч байна гэдгийг харж болох буюу цаашилбал d хэмжээст n ширхэг өгөгдөлийн ихэнхи нь бөмбөрцөгийн гадаргын дагуу байрлаж байна (Зур. 2). Тиймээс ихэнхи тохиолдолд өгөгдөлийн огторгуйг багасгах шаардлагатай болдог (жишээ нь: $d = 3 \rightarrow d = 2$).

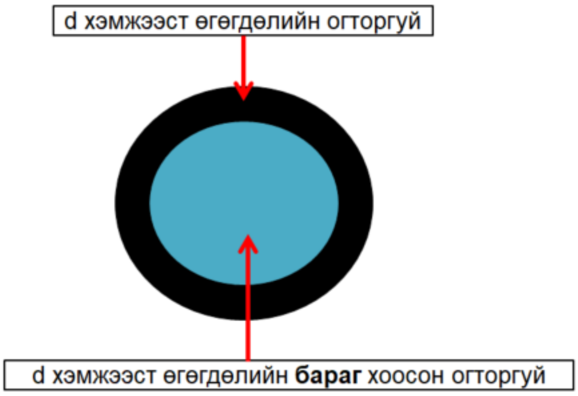


Рис. 50.

26 Магадлалын үндэс

26.1 Олонлог

Олонлог гэдэг нь эрэмбэ, дараалалгүй ялгаатай биетүүдийн цуглуулага. Биетүүдийг нь олонлогийн бүрдүүлэгч элементүүд эсвэл гишүүд гэж нэрлэдэг. Дурын олонлогийг A гэж тэмдэглэе. Харин a нь олонлог A-г бүрдүүлэгч элементүүдийн нэг гэж үзэе. Тэгвэл математикийн хэлээр $a \in A$ гэж илэрхийлдэг. Жишээ 1: Олонлог $A = \{a, b, c, d\}$ нь a, b, c, d бүрдүүлэгч элементүүдийг агуулж байна. Дурын элемент e-г авч үзвэл, энэхүү элементийг олонлог A агуулахгүй байгаа буюу $e \notin A$ гэж илэрхийлдэг. Жишээ 2: 10 хүртэлх тоонд агуулагдаж буй эерэг ба сондгой тооны олонлог нь $B = \{1, 3, 5, 7, 9\}$ бөгөөд үүнийг ихэнхи тохиолдолд доорх байдлаар илэрхийлдэг,

$$B = \{x | \text{x нь 10-аас бага, эерэг ба сондгой too}\}$$

$$= \{x \in \mathbb{Z}^+ | x < 10 \text{ ба x нь сондгой too}\}.$$

26.1.1 Дэд Олонлог

Эерэг ба 10-аас бага тоонуудын олонлогийг авч үзэе;

$$I = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} = \{x \in \mathbb{Z}^+ | x \leq 10\}.$$

Тэгвэл жишээ 2-д дурдсан $B = \{x \in \mathbb{Z}^+ | x < 10 \text{ ба x нь сондгой too}\}$-г I-ын дэд олонлог гэж нэрлэдэг бөгөөд $B \subseteq I$ гэж тэмдэглэдэг. $B \subseteq I$ байх нөхцөлийг математикийн хэлээр

$$\forall x (x \in B \to x \in I). \tag{26.1}$$

Анхаарах зүйл: дурын хоосон биш олонлогт I хоёр дэд олонлог заавал байх ёстой. (1) хоосон олонлог $\varnothing \subseteq I$, (2) Олонлог нь өөрөө, өөрийнхөө дэд олонлог $I \subseteq I$ учир нь тэгшитгэл (25.6) хангагдаж байдаг.

26.1.2 Олонлогийн дүрмүүд

Дурын A ба B олонлогийг авч үзэе. Эдгээр олонлогуудын нэгдлийг $A \cup B$ гэж тэмдэглэдэг. $A \cup B$-г бүрдүүлэгч элементүүд A-д, эсвэл B-д, бүр эсвэл хоёуланд нь зэрэг харъялагдаж байх ёстой;

$$A \cup B = \{x | x \in A \quad \text{or} \quad x \in B\}.$$

Жишээ 3: $A = \{1, 3, 5\}$, ба $B = \{1, 2, 3\}$-ийн нэгдэл нь $A \cup B = \{1, 2, 3, 5\}$. Дурын A ба B олонлогийг авч үзэе. Эдгээр олонлогуудын огтлолцолыг $A \cap B$ гэж тэмдэглэдэг. $A \cap B$-г бүрдүүлэгч элементүүд нь A, B олонлогт хоёуланд нь зэрэг харъялагдаж байх ёстой;

$$A \cap B = \{x | x \in A \quad \text{and} \quad x \in B\}.$$

Жишээ 4: $A = \{1, 3, 5\}$ ба $B = \{1, 2, 3\}$-ийн огтлолцол нь $A \cap B = \{1, 3\}$. Дурын A ба B олонлогийг авч үзэе. Эдгээр олонлогуудын огтлолцол нь хоосон байвал $A \cap B = \varnothing$, A ба B олонлогийг үл давхцах олонлог гэж нэрлэдэг. Жишээ 5: $A = \{1, 3, 5, 7, 9\}$ ба $B = \{2, 4, 6, 8, 10\}$, огтлолцол нь хоосон олонлог $A \cap B = \varnothing$. Жишээ 6: цаашилбал дурын олонлогт агуулагдаж буй элементүүдийн тоог $|\cdot|$ гэж тэмдэглэдэг. Тэгвэл үл давхцах $A \cap B = \varnothing$ олонлогууд $A = \{1, 3, 5, 7, 9\}$ ба $B = \{2, 4, 6, 8, 10\}$ хувьд $|A| = 5, |B| = 5$, $|A \cup B| = 10$ буюу $|A \cup B| = |A| + |B| = 10$ гэж илэрхийлж болно. Давхцаж буй олонлог $A \cap B \neq \varnothing$ сонихрхое. $A = \{1, 3, 5, 7, 10\}$ ба $B = \{2, 4, 6, 8, 10\}$ хувьд $|A| = 5, |B| = 5$, $|A \cup B| = 9$, $|A \cap B| = 1$ буюу $|A \cup B| = |A| + |B| - |A \cap B| = 9$ болно. Тиймээс ерөнхий тэгшитгэл нь

$$|A \cup B| = |A| + |B| - |A \cap B|. \tag{26.2}$$

Дурын A ба B олонлогийг авч үзэе. Эдгээр олонлогуудын ялгааг $A \backslash B$ эсвэл $A - B$ гэж тэмдэглэдэг. $A \backslash B$-г бүрдүүлэгч элементүүд нь A-д агуулагддаг, гэвч B-д агуулагддаггүй байх ёстой, эсвэл эсрэгээрээ.

$$A - B = \{x | x \in A \quad \text{and} \quad x \notin B\},$$

$$B - A = \{x | x \in B \quad \text{and} \quad x \notin A\}.$$

Жишээ 7: $A = \{1, 3, 5\}$, ба $B = \{1, 2, 3\}$. $A - B = \{5\}$ бөгөөд нөгөө талаар $B - A = \{2\}$; $A - B \neq B - A$. Ерөнхий олонлог хэмээх ойлголт байдаг бөгөөд U гэж авч үзэе. Олонлог А-г нэгдмэл болгогч олонлогийг $\bar{A}$ ерөнхий олонлогтой харьцангуй тодорхойлдог ($\bar{A} = U - A$);

$$\bar{A} = \{x \in U | x \notin A\}.$$

Жишээ 8: $U = \{\text{Англи хэлний цагаан толгойн үсгүүд}\}$, $A = \{a, e, i, o, u\}$. Иймээс $\bar{A} = \{b, c, d, f, g, h, j, k, \ldots, z\}$.

26.2 Магадлал

Нэг ширхэг зоосыг авч үзээ. Зоос зөвхөн сүлд эсвэл тоо талаараа буух боломжтой. Тэгвэл сүлд-s, тоо-t гэж тэмдэглэвэл, математикт үүнийг түүврийн огторгуйн олонлог (sample space, Ω) гэж нэрлэдэг $\Omega = \{s, t\}$. Энд s болон t-г олонлогийн гишүүд (members) эсвэл бүрдүүлэгч элементүүд (elements) гэдэг. Түүврийн огторгуйн олонлог-т дэд олонлог/тохиолдолын огторгуй (subset/event space, $\mathbb{F}$) хэмээх ойлголт байдаг. Тохиолдолын огторгуй $\mathbb{F}$ нь түүврийн огторгуйн олонлог Ω-г бүрдүүлэгч элеметүүдийн бүх боломжит сонголтууд $(\{\varnothing\}, \{s\}, \{t\}, \{s, t\})$. Эндээс үзвэл нийт дэд олонлогийн тоо нь 2^n-тэй (n нь олонлог Ω-ын бүрдүүлэгч элементүүдийн тоо) тэнцүү байдаг. Математикийн хэлээр

$$\mathbb{F} = \{\{\varnothing\}, \{s\}, \{t\}, \{s, t\}\} \subseteq \Omega = \{s, t\}; \tag{26.3}$$

$\mathbb{F}$-ын бүрдүүлэгч элементүүд $\{\{\varnothing\}, \{s\}, \{t\}, \{s, t\}\}$-ийг тайлбарлая. Зоосыг $n = 100$-н удаа хаясан гэж үзвэл. Заавал сүлд $\{s\}$, болон тоо $\{t\}$ талаараа буух тохиолдол нь тэнцүү буюу $n_s = 50$ удаа сүлд, $n_t = 50$ удаа тоо гэж авч үзээ. Тэгвэл нийт хаялтын тоо n-г сүлд n_s эсвэл тоо n_t байх тохиолдолд хуваавал

$$P(\{s\}) = P(\{t\}) = \frac{n_s}{n} = \frac{n_t}{n} = 0.5, \tag{26.4}$$

P-г магадлал гэж нэрлэдэг ба s болон t талаараа буух магадлал нь тэнцүү байна. Аль нэг талаараа буухгүй байх тохиолдолыг $\{\varnothing\}$ гэж тэмдэглэсэн бөгөөд магадлалаар илэрхийлбэл $P(\{\varnothing\}) = 0$. Харин сүлд эсвэл тоо (s or t)$\approx$ $\{s, t\}$-ны аль нэг нь заавал буух магадлал нь үргэлж 1-тэй тэнцүү $P(\{s, t\}) = P(s \quad or \quad t) = 1$ юм. Цаашилбал $(\Omega, \mathbb{F}, P)$-ыг Магадлалын Огторгуй гэж нэрлэдэг.

PART

XIII

Хүн-Машин

> Remark (Фёдор Достоевский). хэрэв бүх Дэлхийг эзлэн, ноёрхомоор байвал
> Хамгийн түрүүнд өөрөө, өөрийгөө эзлэн, ноёрх

27 Магадлалын Үндэс: Санамсаргүй тоо

Өмнөх хэсэгт магадлалын талаар ерөнхий ойлголттой болсон. Цаашибал магадлалын талаар ярих бүрд санамсаргүй хувьсагч (random variable) X хэмээх ойлголт гарч ирдэг. Жишээ болгож зургаан талтай (тал бүрийг нь үсгээр тэмдэглэвэл: $f_1, f_2, f_3, f_4, f_5, f_6$) нэг ширхэг шоог сонирхоё.Тэгвэл түүврийн огторгуйн олонлог (sample space)-ийг $\Omega = \{f_1, f_2, f_3, f_4, f_5, f_6\}$ гэж тэмдэглэе. Нийт дэд олонлогийн тоо нь

$$n = 2^6, \mathbb{F} = \{\{\varnothing\}, \{f_1\}, \ldots, \{f_6\}, \ldots, \{f_1, f_2, \ldots, f_6\}\}.$$

Шооны тал бүрд харгалзах магадлал нь тэнцүү $P(\{f_i\}) = \frac{1}{6}, f_i = f_1, f_2, \ldots, f_6$ гэж үзэе. Цаашилбал санамсаргүй хувьсагч хэмээх ойлголтыг тайлбарлахын тулд шооны тал бүрийг тасралтгүй тоон шулуунд орших $\mathrm{X} \in \{10, 20, 30, 40, 60\}$ тоонд харгалзуулая.

$$\mathrm{X} : f_i \to \mathrm{X}(f_i),$$

$\mathrm{X}(f_1) = 10, \mathrm{X}(f_2) = 20, \mathrm{X}(f_3) = 30, \mathrm{X}(f_4) = 40, \mathrm{X}(f_5) = 50, \mathrm{X}(f_6) = 60.$

Магадлалыг зөвхөн тохиолдол (event)-ын хувьд тодорхойлдог учир санамсаргүй хувьсагчийн тохиолдол нь

$$\{\mathrm{X}(f_i) < x\}$$

гэж тодорхойлогддог бөгөөд энд x нь дурын тоо болно. Хэрэв $x = 35$ бол тохиолдол нь

$$\{\mathrm{X}(f_i) < 35\} = \{\mathrm{X}(f_1) = 10, \mathrm{X}(f_2) = 20\}$$

бөгөөд шооны нэгдүгээр эсвэл хоёрдугаар талд нь магадлал $P(\{\mathrm{X}(f_1), \mathrm{X}(f_2)\}) = 2/6$ харгалзуулна гэсэн үг. Дахин жишээ болгож нэгэн зоосыг авч үзээ,

$$\Omega = \{f_1, f_2\}, \mathbb{F} = \{\{\varnothing\}, \{f_1\}, \{f_2\}, \{f_1, f_2\}\}.$$

Цаашилбал зоосны тал бүрийг $\mathrm{X} \in \{0, 1\}$ тоонд харгалзуулая,

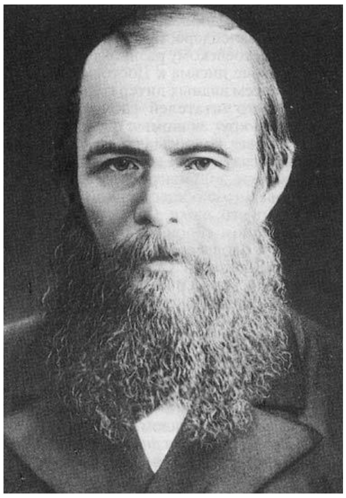


Рис. 51. “Фёдор Михайлович Достоевский (1821 оны 11 сарын 11 [хуучнаар 10 сарын 30] Москва; † 1881 оны 2 сарын 9 [хуучнаар 1 сарын 28] Санкт-Петербург) нь Оросын зохиолч, сэтгэгч, философич, нийтлэлч, сэтгүүлч. 1887 оноос Петербургийн ШУА-ийн сурвалжлагч гишүүн. Оросын реалист утга зохиолд томоохон шинэчлэл хийсэн нь амьд ахуйдаа тухайн цаг үеийнхнээсээ зохих үнэлгээг хүртээгүй. Харин таалал төгссөн хойноо Оросын утга зохиолын сонгодог хэмээн үнэлэгдсэн төдийгүй зохиол туурвил нь дэлхийн утга зохиолд, тэр дундаа Нобелийн шагналт олон зохиолчийн бүтээлд жинтэй нөлөөлсөн бөгөөд экзистенциализм, фрейдизмийг бүрэлдэн тогтоход ч зохих нөлөө үзүүлсэн гэж үздэг.”

$$\mathrm{X}(f) = \begin{cases} 1, & f = f_1 \quad \text{буюу f нь сүлд үед.} \\ 0, & f = f_0 \quad \text{буюу f нь тоо үед.} \end{cases} \tag{27.1}$$

буюу хэрэв зоосыг 100 удаа хаявал 100 ширхэг санамсаргүй хувьсагч $\{0, 1\}$-үүдийн цуглуулгыг олж авах болно. Дүгнэж хэлбэл санамсаргүй хувьсагч (random variable) X нь түүврийн огторгуйд Ω харъялагдах үр дүнгүүдийг бодит тоонд харгалзуулдаг буюу буулгадаг. Дээрх жишээнүүд нь түүврийн огторгуйн Ω үр дүнгүүд нь тасралттай үед тодорхойлогдсон билээ. Гэвч бидний амьдарч буй хүрээлэн буй орчинд явагддаг физик үзэгдлүүд нь тасралтгүй хувьсагчуудаар тодорхойлогддог. Үүний нэгэн жишээ бол тодорхой цаг, хугацааны завсар дахь машины хурд $v(t)$ болон зам $s(t)$ билээ. Учир нь машин a-аас b цэгд очихын тулд үргэлж тэгээс ялгаатай хурдтай байх хэрэгтэй. Тиймээс тасралтгүй санамсаргүй хувьсагч X-ийн тодорхой утгаас x бага буюу тэнцүү байх магадлалыг cumulative түгэлтийн функц, cumulative distribution function (CDF) гэж нэрлэдэг,

$$P(\{\mathrm{X} \leq x\}) = F_{\mathrm{X}}(x). \tag{27.2}$$

Дээр дурдсан нэг ширхэг зоосны хувьд CDF-ийг тооцож үзэе. $x = 10$ үед тэгшитгэл (27.1)-ийн дагуу $\mathrm{X}(f_1) = 1 \leq x, \mathrm{X}(f_0) = 0 \leq x$ нөхцөлүүд хангагддаг бол

$$F_{\mathrm{X}}(10) = P(\{\mathrm{X}(f) \leq x\}) = P(\{\mathrm{X}(f_0), \mathrm{X}(f_1)\}) = 1,$$

$x = 0.4$ үед тэгшитгэл (1)-ийн дагуу $\mathrm{X}(f_0) = 0 \leq x$ нөхцөл хангагддаг бол

$$F_{\mathrm{X}}(0.4) = P(\{\mathrm{X}(f) \leq x\}) = P(\{\mathrm{X}(f_0)\}) = 1/2.$$

Ихэнхи тохиолдолд $F_{\mathrm{X}}(x) = F(x)$ гэж тэмдэглэдэг. CDF нь санамсаргүй хувьсагч x-ээс хамаардаг учир уламжлалыг нь тодорхойлох боломжтой;

$$\frac{dF(x)}{dx} = p(x), \tag{27.3}$$

үүнийг магадлалын нягтын функц (probability density function, pdf) гэж нэрлэдэг төдийгүй $p(x) \geq 0$ ба $F(+\infty) = 1, F(-\infty) = 0$ нөхцөлүүд заавал хангагдах ёстой. Тэгшитгэл (27.3)-ыг $[-\infty, +\infty]$ мужид интегралчилбал

$$F(x) = \int_{-\infty}^{x} p(\epsilon) d\epsilon, \int_{-\infty}^{+\infty} p(x) dx = 1. \tag{27.4}$$

Цаашилбал

$$\int_a^b p(x) dx = P(\{a \leq \mathrm{X} \leq b\}) = F(b) - F(a). \tag{27.5}$$

Инженерийн хэрэглээнд түгээмэл хэрэглэгддэг нь нэгэн төрлийн түгэлтийн функц $\mathbb{U}([a, b])$ бөгөөд pdf нь доорх байдлаар тодорхойлогддог,

$$p_{\mathbb{U}} = \frac{1}{b-a}$$

Өөр нэгэн түгээмэл хэрэглэгддэг нь нормал түгэлтийн функц $\mathbb{N}(\mu, \sigma)$ бөгөөд pdf нь доорх байдлаар тодорхойлогддог,

$$p_{\mathbb{N}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right\}.$$

27.1 Варианс болон хазайлт

Санамсаргүй хувьсагчийн X дундаж утга буюу expected value $\mathbb{E}$ нь утгын мужид Ω доорх байдлаар тодорхойлогддог,

$$\mathbb{E}[\mathrm{X}] = \langle \mathrm{X} \rangle = \int_{\Omega} xp(x)dx. \tag{27.6}$$

Тасралттай хувьсагчийн хувьд дундаж утга нь

$$\mathbb{E}(x) = \sum_i x_i P(x_i). \tag{27.7}$$

Цаашилбал санамсаргүй хувьсагчуудын X, Y хувьд дурын тооны хувьд a, b доорх шинж чанар хангадаг

$$\mathbb{E}[a\mathrm{X} + b\mathrm{Y}] = a\mathbb{E}[\mathrm{X}] + b\mathbb{E}[\mathrm{Y}]. \tag{27.8}$$

Санамсаргүй хувьсагчийн X варианс буюу дундаж утгаас хазайх хазайлт нь

$$Var[\mathrm{X}] = \sigma^2[\mathrm{X}] = \mathbb{E}[(\mathrm{X} - \mathbb{E}[\mathrm{X}])^2]. \tag{27.9}$$

Хэрэв дурын 2 санамсаргүй хувьсагчууд X, Y нь хоорондоо хамааралгүй байвал эдгээрийн үржвэрийн expected value нь тус бүрийн үржвэрээр тодорхойлогдоно,

$$\mathbb{E}[\mathrm{XY}] = \mathbb{E}[\mathrm{X}]\mathbb{E}[\mathrm{Y}]. \tag{27.10}$$

Мөн түүнчилэн варийнсын хувьд шугаман шинж чанар хадгалагддаг

$$\sigma^2\left[\sum_i a_i \mathrm{Y}_i\right] = \sum_i a_i^2 \sigma^2[\mathrm{Y}_i]. \tag{27.11}$$

28 Монте Карло Интегралчилал

Монте Карло Интегралчилалын аргын талаар тайлбарлахын өмнө, π тоог хэрхэн энэхүү аргыг ашиглан тооцох талаар сонирхоё. Бид r радиустай тойргийн талбай нь πr^2 гэдгийг мэднэ. Тойргийг ашиглан π тоог тооцохын тулд тооцоогоо бяцхан хялбрчилан $r = 1$ бөгөөд тойргийн дөрөв хуваасны нэгийг авч үзэе. Энэ тохиолдолд тойргийн 1/4-ийн талбай нь $\pi/4$-тэй тэнцүү болно. Математикийн хэллэгээр тойргийн 1/4 хэсгийн талбай нь $f : [0, 1]^2 \to \mathbb{R}$ функцын интегралтай тэнцүү буюу функц нь доорх шинж чанартай,

$$f(x,y) = \begin{cases} 1, & \text{if} \quad \sqrt{x^2+y^2} \le 1 \\ 0, & \text{else.} \end{cases} . \tag{28.1}$$

Функцын $\int_0^1 f(x,y)dxdy$ интеграл нь доор үзүүлсэн зурагны усан сангийн тал-

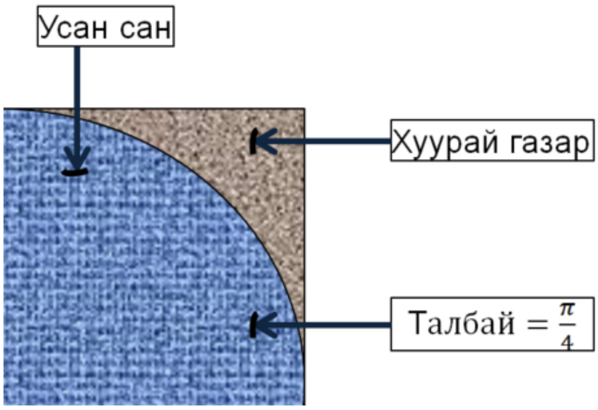


Рис. 52. Pond-Ground

байтай тэнцүү буюу үгээр илэрхийлбэл, бид нүдээ аниж байгаад чулууг дээрх дөрвөлжин дүрс үрүү шидлээ гэж бодоё. Хэрэв чулуу усанд унавал 1, харин газарт унавал тоолохгүй буюу 0 гэж үзвэл, нийт усанд унасан чулуу нь $\pi/4$-тэй тэнцүү байх болно. Эдгээр шидсэн чулууг түүврийн санамсаргүй цэг (random sampling point)-үүд гэж хэлдэг. Цаашилбал өгөгдсөн d хэмжээст функцын хувьд, $f : \mathbb{R}^d \to \mathbb{R}$, интеграл нь

$$I = |\Omega|\langle f \rangle \tag{28.2}$$

$\langle f \rangle = \int_\Omega f(\vec{x})p(\vec{x})d\vec{x}$ бөгөөд $p(\vec{x}) = 1/|\Omega|$ нэгэн төрлийн функц үед

$$\langle f \rangle = \frac{1}{|\Omega|}\int_\Omega f(\vec{x})d\vec{x}. \tag{28.3}$$

Өмнөх хэсэгт дурдсанчилан $\langle f \rangle$-ийг тооцохын тулд f-ийг n^d цэгүүдэд авч үзэх шаардлагатай болдог буюу үүнийг curse of dimensionality гэж хэлдэг гэж үзсэн. Үүнийг даван туулахын тулд M ширхэг түүврийн санамсаргүй цэг (random sampling point)-үүдийг нэгэн төрлийн функцээс түүвэрлэн аваад, дундаж утгыг нь доорх байдлаар тооцдог

$$\langle f \rangle \approx \langle f \rangle_M = \frac{1}{M}\sum_{i=1}^{M} f(\vec{x}_i). \tag{28.4}$$

Үүнийг Монте Карло Интегралчилал-ын арга гэж хэлдэг. Энгийн жишээ болгон $d = 1$ нэг хэмжээст функцын дундаж утгыг , $|\Omega| = b - a$, $M = 4$ тохиолдолд зургаар илэрхийлэе.

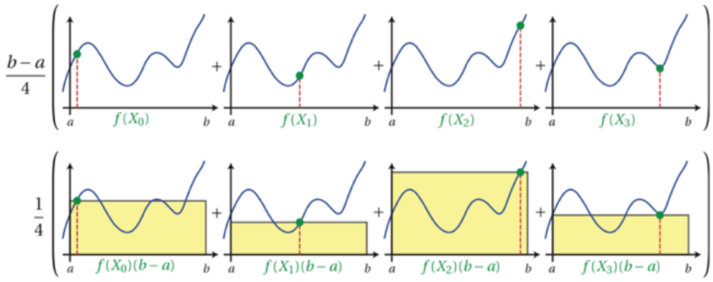


Рис. 53. (Source: computational methods ETH)

Анхаарах зүйл бол бидний тооцож буй функц $\langle f \rangle_M$ нь түүврийн санамсаргүй цэг-үүдээс хамаарч буй учир $\langle f \rangle_M$ мөн санамсаргүй дурын утга болж хувирдаг. Иймээс энэнүү функцын expected value-г нь авч үзэх шаардлагатай болдог,

$$\mathbb{E}[\langle f \rangle_M] = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M} f(\vec{x}_i)\right] = \frac{1}{M}\sum_{i=1}^{M}\langle f \rangle = \langle f \rangle, \tag{28.5}$$

энэ нь бодит утгатай тэнцүү байдаг. Монте Карло Интегралчилал-ын алдаа нь

$$\begin{aligned} e_M &= \sqrt{Var[\langle f \rangle_M - \langle f \rangle]}, \quad \mathbb{E}[\langle f \rangle_M] = \langle f \rangle \\ e_M &= \sqrt{Var[\langle f \rangle_M]}, \quad e_M = \sqrt{\frac{Var[f]}{M}} = \mathbb{O}(M^{-1/2}). \end{aligned} \tag{28.6}$$

Эндээс үзвэл Монте Карло Интегралчилал-ын аргын алдаа нь өмнөх хэсэгт дурдсан интегралыг тоон аргаар ойролцоолох аргуудаас бага байна гэдэг нь харагдаж байна.

29 Урвуу Хувиргалтын Түүвэр (Inverse Transform Sampling)

Магадлалын нягтын функц pdf $p(x)$ ба CDF $F_{\mathrm{X}}(x)$-ын дагуу тарахлттай санамсаргүй хувьсагч X-ийг авч үзэе. pdf $p(x)$-ээс түүврийн цэгүүдийг үүсгэхийн тулд урвуу хувиргалтын түүврийн аргыг ашиглаж болно. Энэхүү арга нь нэгэн төрлийн түгэлтийн $U([a, b])$ дагуу тархалттай санамсаргүй хувьсагчуудын түүврүүдийг $u^{(i)} (i = 1, 2, \ldots, N)$ ашигладаг. Иймээс бидэнд санамсаргүй хувьсагчууд болох X болон U-ийг холбосон хамаарлын хувиргалтын тэгшитгэл хэрэгтэй бөгөөд доорх байдлаар тодорхойлоё,

$$x = g(u). \tag{29.1}$$

Нэгэн төрлийн түгэлттэй санамсаргүй хувьсагчийн хувьд U, pdf $p_U(u) = 1, u \in [0, 1]$ болон CDF нь

$$F_U(u) = \int_0^u p_U(t)dt = u. \tag{29.2}$$

Цаашилбал дурын CDF $F_{\mathrm{X}}(x) \in [0,1]$ хувьд бид

$$\begin{aligned} F_{\mathrm{X}}(x) &= u, \\ x &= F_{\mathrm{X}}^{-1}(u). \end{aligned} \tag{29.3}$$

хэмээн тодорхойлох боломжтой. Сүүлийн тэгшитгэл нь pdf $p_{\mathrm{X}}(x)$-ээс түүврийг үүсгэх боломжийг олгож байна,

$$x^{(i)} = F_{\mathrm{X}}^{-1}(u^{(i)}). \tag{29.4}$$

Хялбар жишээ болгон экспоненциал түгэлтийн функцээс түүврийг үүсгэе $x^{(i)}$,

$$p(x) = \lambda e^{-\lambda x}, x > 0. \tag{29.5}$$

Тэгвэл тэгшитгэл (29.2)-ын дагуу CDF нь

$$F(x) = \int_0^x p(x)dx = 1 - e^{-\lambda x}, \tag{29.6}$$

мөн тэгшитгэл (29.3)-ын $u = F(x)$ дагуу

$$\begin{aligned} 1 - e^{-\lambda x} &= u, \\ x &= -\frac{1}{\lambda}\ln(1-u). \end{aligned} \tag{29.7}$$

Сүүлийн тэгшитгэлийн x нь экспоненциал түгэлттэй бөгөөд иймээс энэхүү түгэлт (29.5)-ээс түүврүүдийг $x^{(i)}$ үүсгэх боломж олгодог,

$$x^{(i)} = -\frac{1}{\lambda}\ln\Big(1 - u^{(i)}\Big). \tag{29.8}$$

30 Rejection Sampling

Санамсаргүй тоо X-ийг үүсгэх өөр нэгэн арга бол Вон Нёманны санал болгосон Rejection Sampling юм. Энэхүү арга нь санамсаргүй тоог магадлалын түгэлтийн функц $h(x)$-ээс үүсгэх бөгөөд $p(x)$-ийн дээд зааг байдаг,

$$p(x) < \lambda h(x), \tag{30.1}$$

энд $\lambda \in \mathbb{R}$ дурын тоо, харин $h(x)$ нь ихэнхи тохиолдолд нэгэн төрлийн түгэлтийн функц байдаг. Алгоритм:

- Санамсагүй түүвэр x-ийг магадлалын түгэлтийн функц $h(x)$-ээс үүсгэнэ.
- Санамсаргүй нэгэн төрлийн түүвэр u-ийг нэгэн төрлийн магадлалын түгэлтийн функцын $[0,1]$ завсраас үүсгэнэ.
- Зөвшөөр (accept) x-ийг хэрэв $u < \frac{p(x)}{\lambda h(x)}$ бол, эс бөгөөд татгалз (reject).

- Хангалттай санамсаргүй түүвэртэй x болох хүртэл дээрх алхмыг давтан гүйцэтгэнэ.

Их хэмжээст $d > 3$ огторгуйн хувьд curse of dimensional үүсдэг учир ихэнхи санамсаргүй түүврүүд x нь татгалзах хэсэгт оршидог. Эсвэл түгэлтийн функцын завсар нь маш бага үед санамсаргүй түүврүүд x нь татгалз хэсэгт оршидог. Иймээс их хэмээст түгэлт болон бага завсрын түгэлтийн функцын хувьд өөр нэгэн арга хэрэгтэй болж байна.

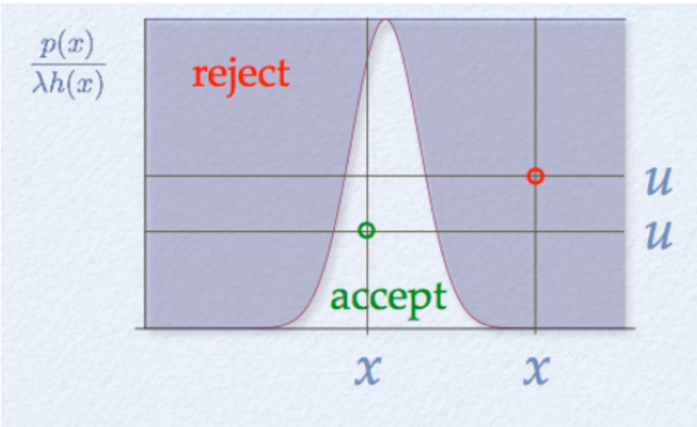


Рис. 54. Rejection Sampling (Source: computational methods ETH).

31 Importance Sampling

Importance Sampling нь бидэнд дээр дурдсан асуудлыг даван гарахад тусалдаг бөгөөд санамсаргүй хувьсагчийг нэгэн төрлийн түгэлтийн функцээс түүвэрлэхээс илүү магадлалын түгэлтээс $w(x)$ түүвэрлэх боломжийг олгодог. Өөр үгээр илэрхийлбэл бид бүхэн магадлалын түгэлтийн функцыг $p(x)$ нормалчиладаг $p(x)/w(x)$ тэгвэл

$$\langle f \rangle_p = \int_a^b f(x) \frac{p(x)}{w(x)} w(x) dx \approx \frac{1}{M} \sum_{i=1}^{M} f(x_i) \frac{p(x_i)}{w(x_i)}, \qquad (31.1)$$

энд x_i бүрийг $w(x)$-ээс түүвэрлэсэн.

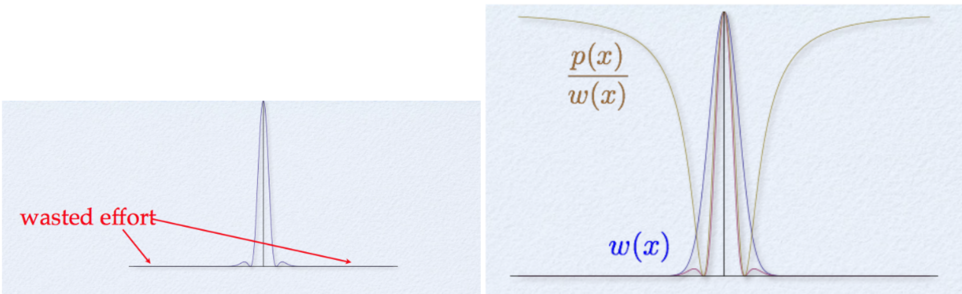


Рис. 55. Importance Sampling (Source: computational methods ETH).