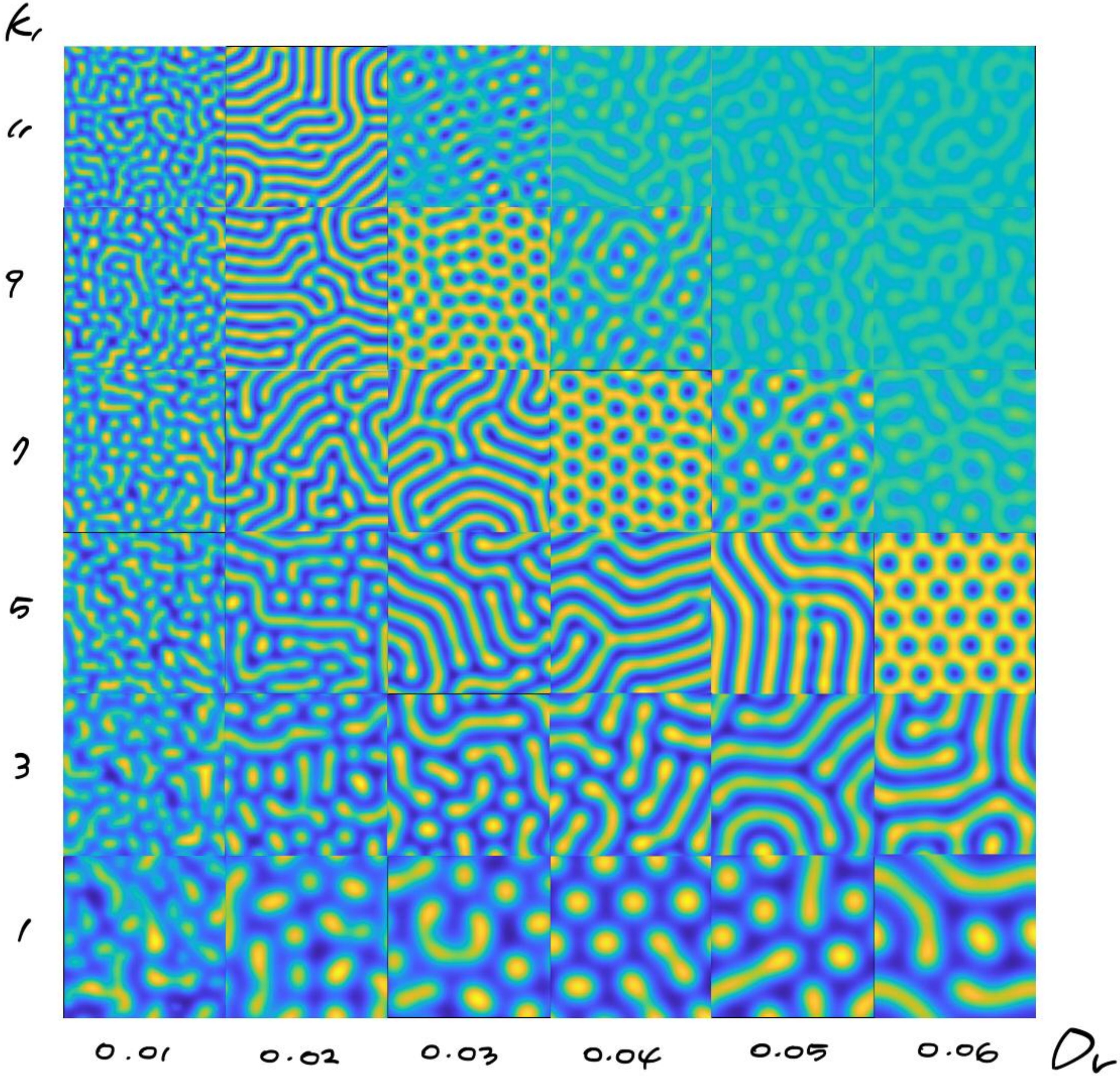
Gradient Descent

2017010698
수학과 오서영



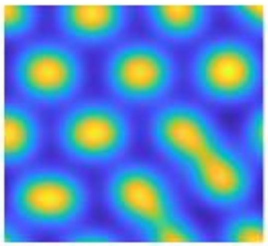
k_r
11
9
7
5
3
1
0.01
0.02
0.03
0.04
0.05
0.06
D_r

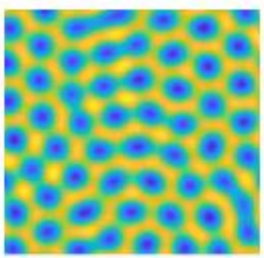
```
% set the parameters (time discretization)
dt=0.1*h^2;
maxit=80000;
nn=1000;

% set the initial condition
u=ubar+0.1*(2*rand(nx+2,ny+2)-1);
v=vbar+0.1*(2*rand(nx+2,ny+2)-1);
nu=u; nv=v;
```

random perturbation

input X — output Y

 — ①

 — ②

 — ③

"random" perturbation

↓

각각 1000개의 data

1 ~ 3

$$\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \cdots, (x^{(3000)}, y^{(3000)})\}$$

image

training test

$$x \rightarrow z = \underbrace{wx+b}_{\text{training}} \rightarrow g(z) \rightarrow \hat{y}$$

for training w, b (parameters)

we need to define "cost function"

given $\{(x^{(1)}, y^{(1)}), \cdots, (x^{(3000)}, y^{(3000)})\}$

want $\underline{\hat{y}^{(i)}} \cong y^{(i)}$

(i) 번째 training sample의
예상치

Loss (error) function

: 예측값 ($\hat{y}$)와 실제값 (y)에 대한 오차

ex) $0.5 \| \hat{y} - y \|^2$

Cost function

: Loss function의 평균

$$" J = \frac{1}{\underset{3000}{m}} \sum_{i=1}^{m} \underbrace{L(\hat{y}^{(i)}, y^{(i)})}_{\text{Loss}}$$

"training"

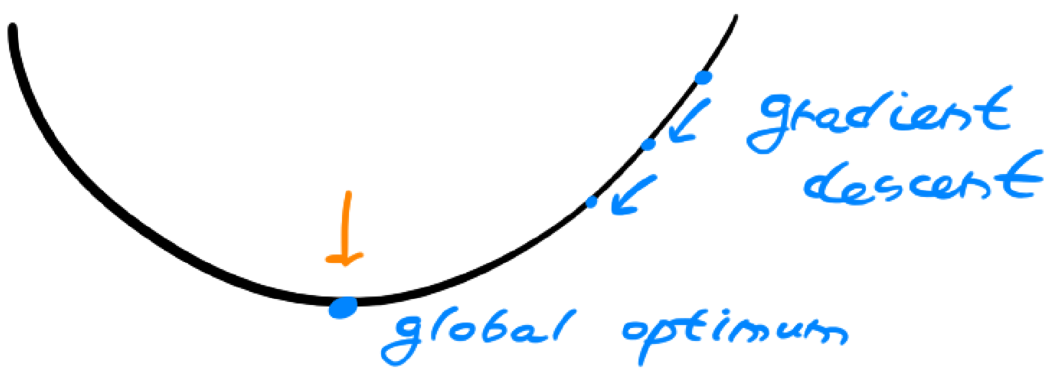
⇒ find w, b

+ minimize J

Gradient Descent

: want to find w, b
that minimize J

→ J(w, b) : convex



$f(x)$: differentiable function

x_1 : initial parameter

iteratively moving x_1 to lower value of $f(x)$

how?

방향

속력

Taylor Expansion

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + O(\|x-a\|^2)$$

assume

$$a = x_1 \ , \ x = x_1 + hu$$

unit direction vector

$$f(x_1+hu) = f(x_1) + hf'(x_1)u + h^2 O(1)$$

$$\Rightarrow f(x_1+hu) - f(x_1) \approx hf'(x_1)u$$

$$\text{"} u^* = \arg\min_{u} \{ f(x_1+hu) - f(x_1) \}$$

$$= \arg\min_{u} hf'(x_1)u = -\frac{f'(x_1)}{\|f'(x_1)\|} \text{"}$$

$$x_{t+1} \underset{\text{update}}{\leftarrow} x_t + h u_1^* = x_t - h \frac{f'(x_1)}{\|f'(x_1)\|}$$