卷积神经网络
Convolutional neural network（CNN）

本节课内容

1. **图像分类**

    1.1 线性分类器

    1.2 损失函数

    1.3 梯度下降

    1.4 反向传播

2. **卷积神经网络**

    2.1 卷积层

    2.2 池化层

    2.3 激活函数

    2.4 全连接层

3. **经典CNN（AlexNet）**

1. 图像分类

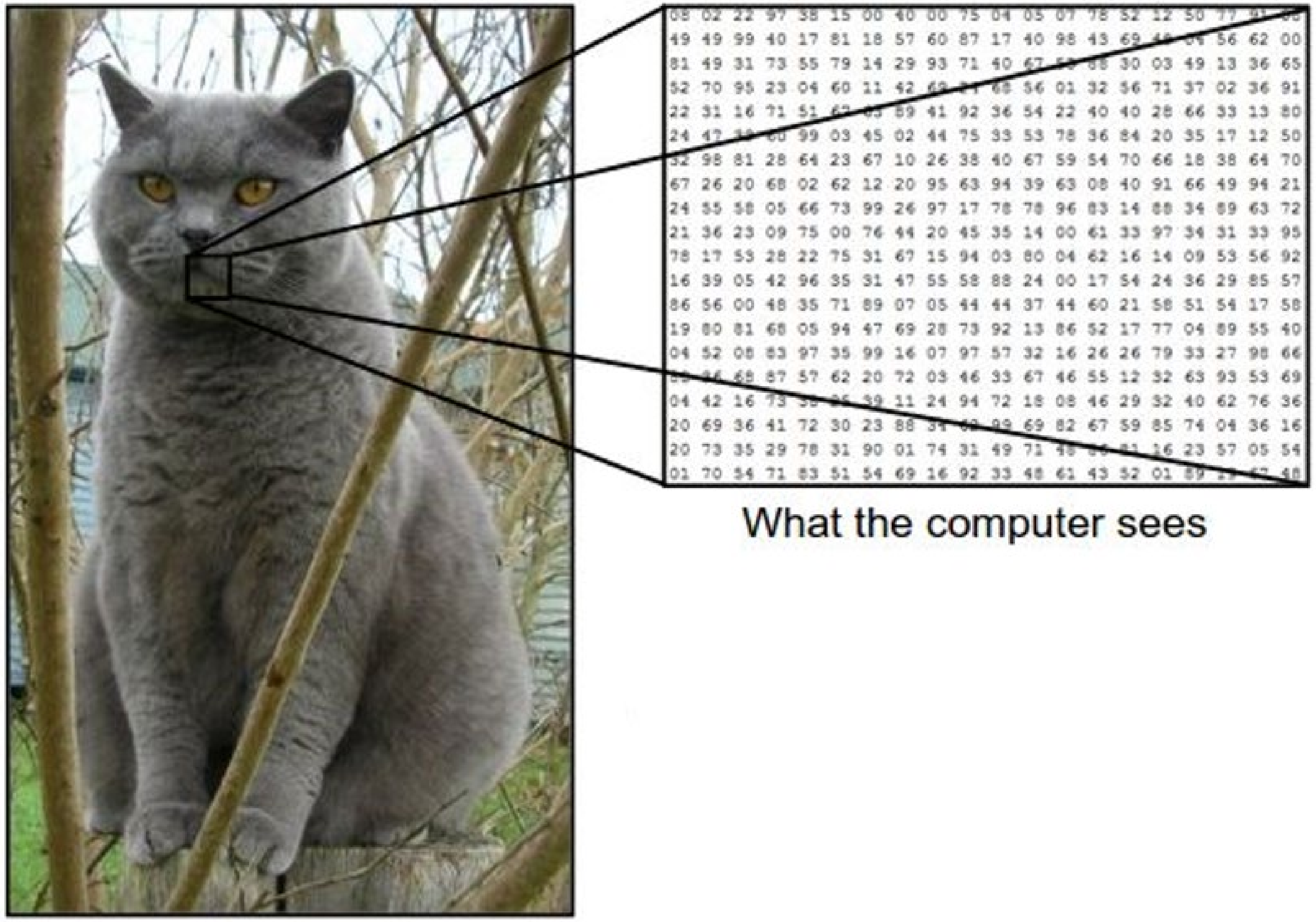
图像分类——计算机视觉的核心问题

假设给定一组离散标签:dog, car, cow, automobile …

cat

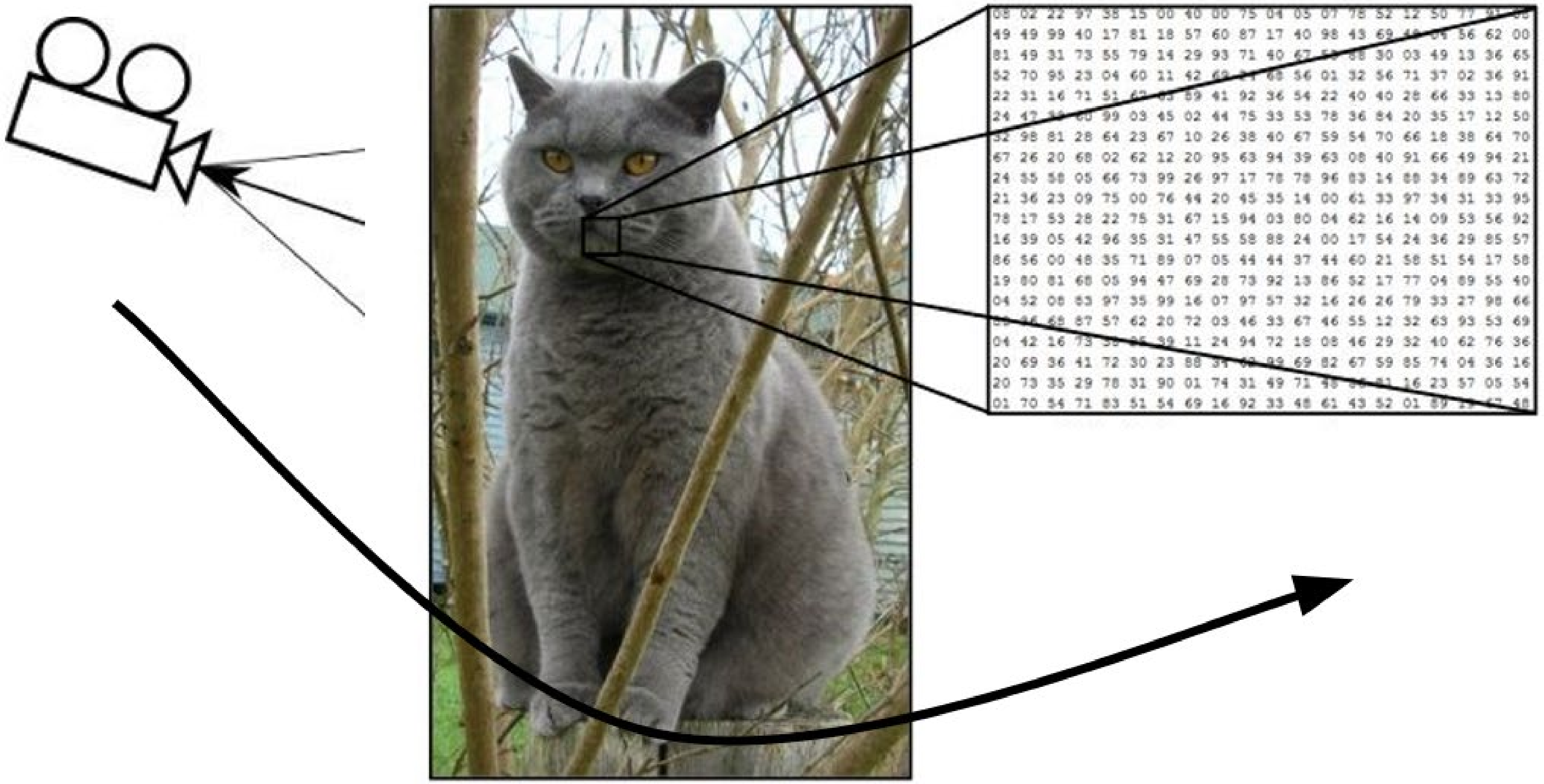
图片的语义鸿沟

- 一个RGB图片可以被表示成3D的数字阵列
  (3维张量tensor)
- 取值范围为[0,255]
- E.g. 200 x 100 x 3 （3表示三个通道）



What the computer sees

挑战——视角变化（Viewpoint Variation）



挑战——光照条件（Illumination）

挑战——形变（Deformation）

猫的行为艺术

挑战——遮挡（Occlusion）

You cannot see me!

挑战——背景干扰（Background clutter）

挑战——类间细分（Intraclass variation）

上一节课回顾

- 数字图像处理
  - 二值图、灰度图、彩色图
  - 数字图像采样（Image Sampling）
  - 数字图像操作（Image Augmentation）
  - 数字图像滤波（Image Filter）

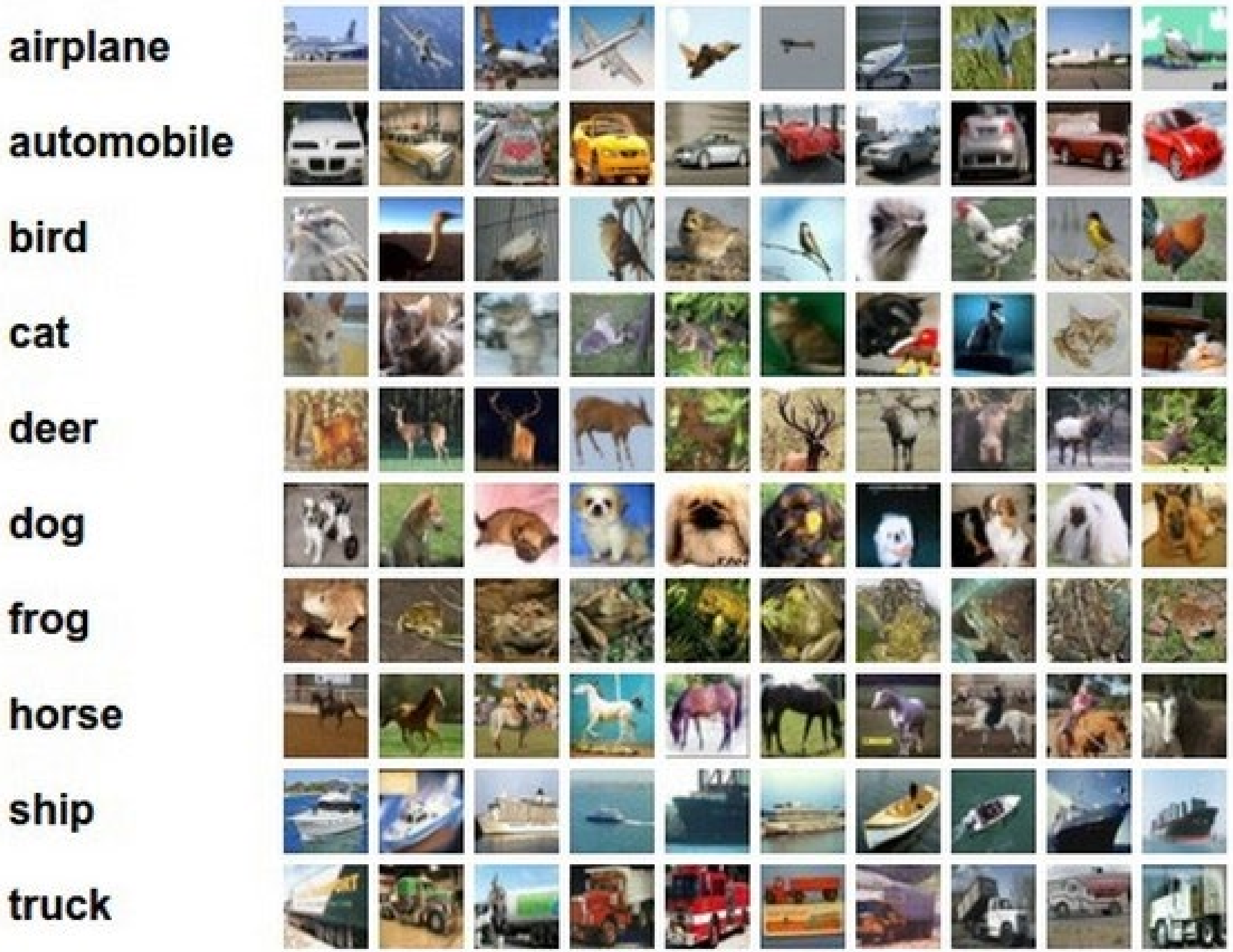
- 图像识别的挑战

- 参考文献来源
  - 期刊：PAMI，IJCV，TIP，CVIU
  - 会议：CVPR，ICCV，ECCV，ICML，ICLR，AAAI，NeurIPS

Anaconda

- https://www.anaconda.com/distribution/
- Anaconda中包含Python、科学包及依赖库
- 学习使用Jupyter（是一个交互式笔记本，Kaggle常用工具）

1.1 线性图像分类器



Example dataset: CIFAR-10

10 labels

50,000 training images

10,000 test images

each image is 32x32x3.

- 图像分类器

input image

array: [32x32x3]
共3072像素

输入图像 **参数**

$f(x, W)$

分类器

输出[10*1]
表示10个类别分数

CIFAR-10
10 labels

- 线性图像分类器

input image

array：[32x32x3]

共3072像素

线性分类器

$$\boxed{f(x, W)} = \boxed{W}\boxed{x}\ \boxed{(+b)}$$

3072*1

10*1

10*1

10*3072

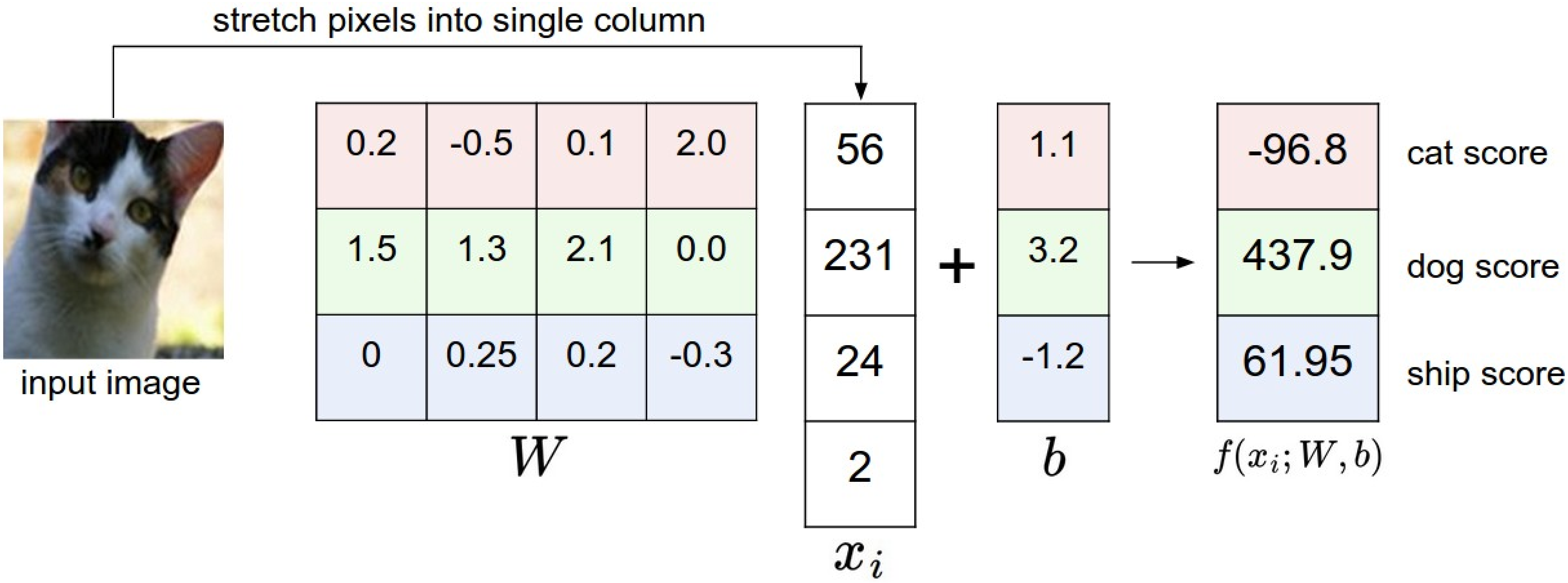
“偏置”（bias）

10个数

表示类别分数

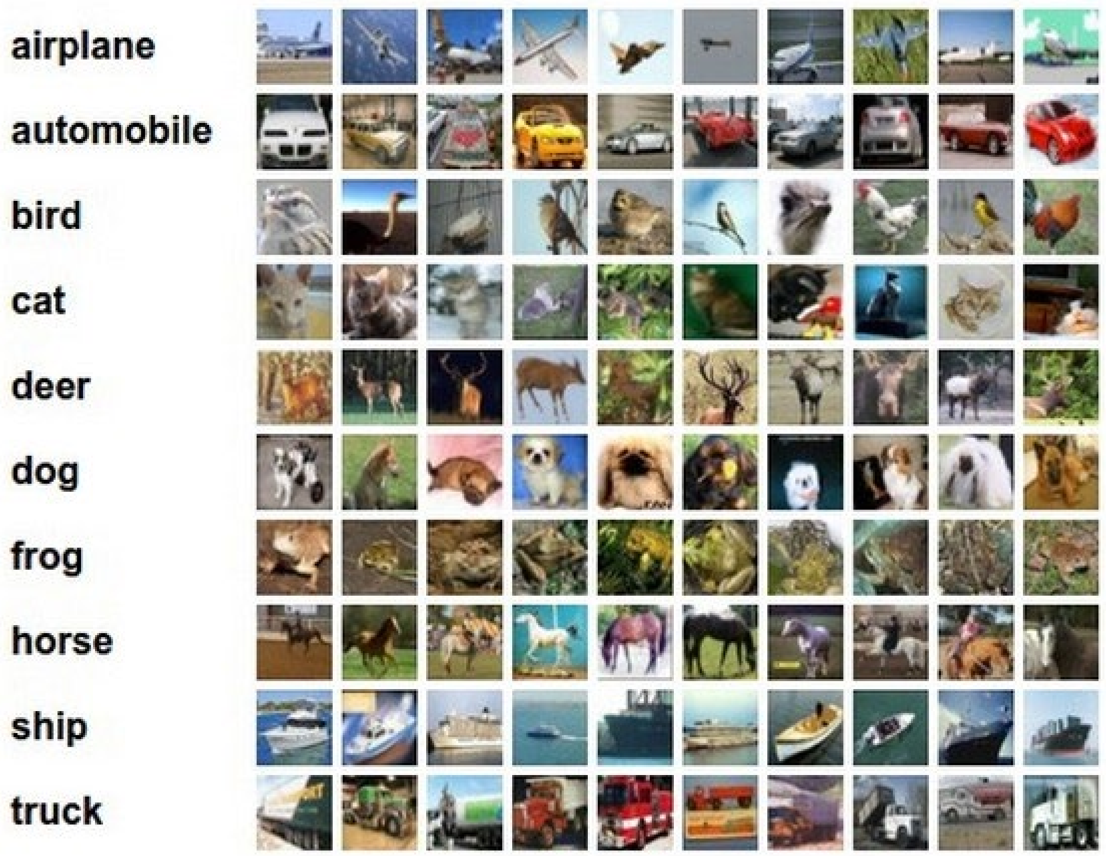
“权重”（weights）

- 线性分类器



- 示例：输入图像4像素，有3类（cat/dog/ship）

- 线性分类器



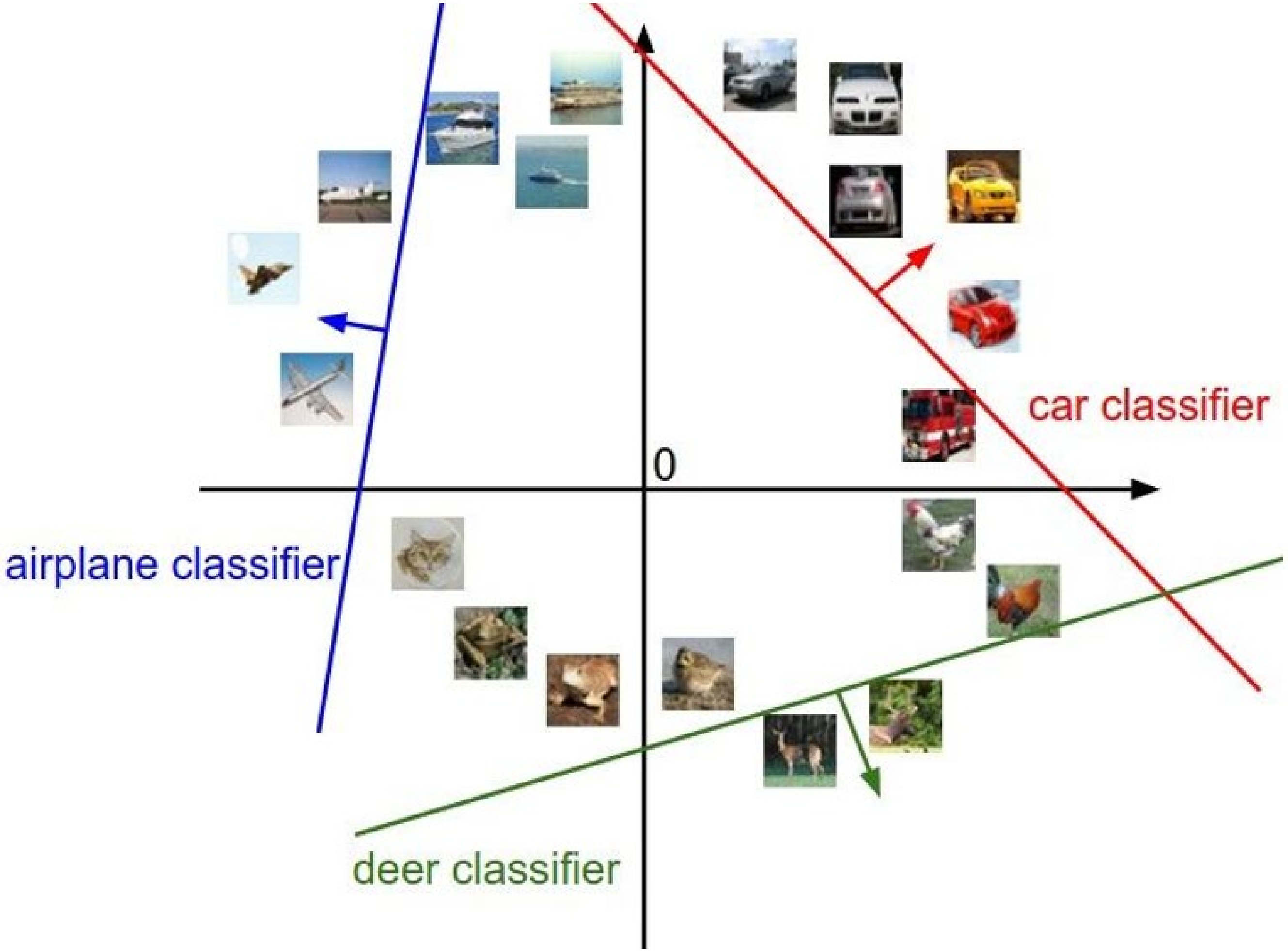
线性分类器

$$f(x_i;\boxed{W},b) = Wx_i + b$$

线性分类器的权重示例

Trained on CIFAR-10

- 线性分类器



$$f(x_i, W, b) = Wx_i + b$$

[32x32x3]

- 线性分类器

score function: $f(x_i;W,b)=Wx_i+b$



随机设定权重W
3张示例图的类别得分

airplane	-3.45	-0.51	3.42
automobile	-8.87	$\boxed{6.04}$	4.64
bird	0.09	5.31	2.65
cat	**2.9**	-4.22	5.1
deer	4.48	-4.19	2.64
dog	$\boxed{8.02}$	3.58	5.55
frog	3.78	4.49	**-4.34**
horse	1.06	-4.37	-1.5
ship	-0.36	-2.09	-4.79
truck	-0.72	-2.93	$\boxed{6.14}$

输入为*automobile* 输出*automobile*得分最高 √

输入为*cat* 输出*dog*得分最高 ×

输入为*frog* 输出*truck*得分最高 ×

分类结果：

	×	√	×
airplane	-3.45	-0.51	3.42
automobile	-8.87	**6.04**	4.64
bird	0.09	5.31	2.65
cat	**2.9**	-4.22	5.1
deer	4.48	-4.19	2.64
dog	8.02	3.58	5.55
frog	3.78	4.49	**-4.34**
horse	1.06	-4.37	-1.5
ship	-0.36	-2.09	-4.79
truck	-0.72	-2.93	6.14

给定随机权重W，分类结果差

- 解决方案：

  定义损失函数，通过优化权重W来最小化损失函数，以找到最好的分类器，达到最好的分类结果。

1.2 损失函数（loss function）

定义loss function：L_i → 衡量分类结果与真实值的不一致程度

1	cat	**3.2**	1.3	2.2
2	car	5.1	**4.9**	2.5
3	frog	-1.7	2.0	**-3.1**

Multiclass loss:

给定一个训练样本（x_i，y_i）

x_i为输入图像

y_i为输出label∈[1, 2, 3]

记 $s_i = f(x_i, W) = Wx_i$

$$L_i = \sum_{j \neq y_i} max(0, s_j - s_{y_i} + 1)$$

示例：训练样本3张，有3类（cat 1/car 2/frog 3）

损失函数计算

定义loss function：L_i → 衡量分类结果与真实值的不一致程度



1	cat	**3.2**	1.3	2.2
2	car	5.1	**4.9**	2.5
3	frog	-1.7	2.0	**-3.1**
loss:		2.9		

Multiclass loss:

$$L_1 = \sum_{j \neq y_1} \max(0, s_j - s_{y_1} + 1)$$

$=\max(0,\ 5.1 - 3.2 + 1)$

$+\max(0,\ -1.7 - 3.2 + 1)$

$=\max(0,\ 2.9)+\max(0,\ -3.9)$

$=2.9+0$

$=2.9$

损失函数计算

定义loss function：L_i → 衡量分类结果与真实值的不一致程度

				
1	cat	**3.2**	1.3	2.2
2	car	5.1	**4.9**	2.5
3	frog	-1.7	2.0	**-3.1**
loss:		2.9	0	

Multiclass loss:

$$L_2 = \sum_{j \neq y_2} \max(0, s_j - s_{y_2} + 1)$$

$=\max(0,\ 1.3 - 4.9 + 1)$

$+\max(0,\ 2.0 - 4.9 + 1)$

$=\max(0,\ -2.6)+\max(0,\ -1.9)$

$=0+0$

$=0$

损失函数计算

定义loss function: L_i → 衡量分类结果与真实值的不一致程度

1	cat	**3.2**	1.3	2.2
2	car	5.1	**4.9**	2.5
3	frog	-1.7	2.0	**-3.1**
loss:		2.9	0	12.9

Multiclass loss:

$$L_3 = \sum_{j \neq y_3} \max(0, s_j - s_{y_3} + 1)$$

$=\max(0,\ 2.2-(-3.1)+1)$

$+\max(0,\ 2.5-(-3.1)+1)$

$=\max(0,\ 6.3)+\max(0,\ 6.6)$

$=6.3+6.6$

$=12.9$

损失函数计算

定义loss function: L_i → 衡量分类结果与真实值的不一致程度

1	cat	**3.2**	1.3	2.2
2	car	5.1	**4.9**	2.5
3	frog	-1.7	2.0	**-3.1**
loss:		2.9	0	12.9

Multiclass loss:

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

总的训练损失：

$$L = \frac{1}{N}\sum_{i=1}^{N} L_i$$

=(2.9+0+12.9)/3

=5.3

分类结果：



	×	√	×
airplane	-3.45	-0.51	3.42
automobile	-8.87	**6.04**	4.64
bird	0.09	5.31	2.65
cat	**2.9**	-4.22	5.1
deer	4.48	-4.19	2.64
dog	8.02	3.58	5.55
frog	3.78	4.49	**-4.34**
horse	1.06	-4.37	-1.5
ship	-0.36	-2.09	-4.79
truck	-0.72	-2.93	6.14

给定随机权重W，分类结果差

- 解决方案：

  定义损失函数，通过优化权重W来最小化损失函数，以找到最好的分类器，达到最好的分类结果。

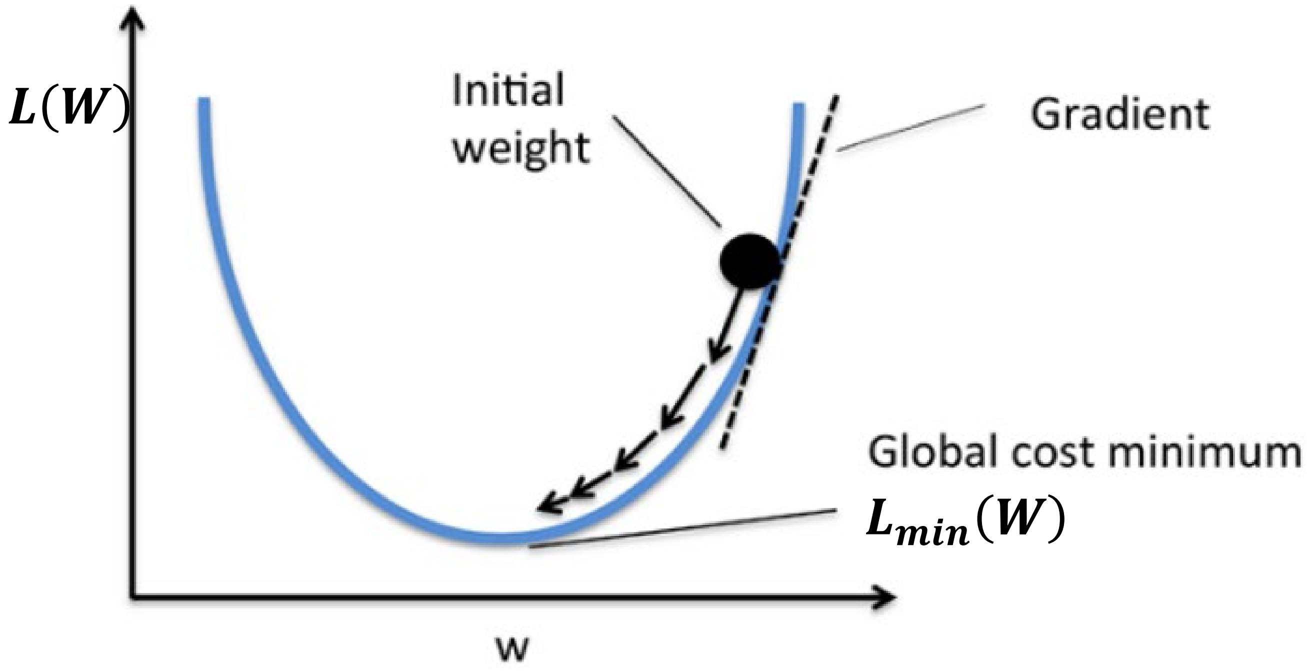
梯度下降

1.3 梯度下降

- **问题**

如何优化调整参数W,b，以达到损失函数最小化的目的?

解决方案——梯度下降用于更新权重W，以及偏置b。



1. 初始化参数，计算损失L;
2. 梯度下降法每次迭代：

? ?

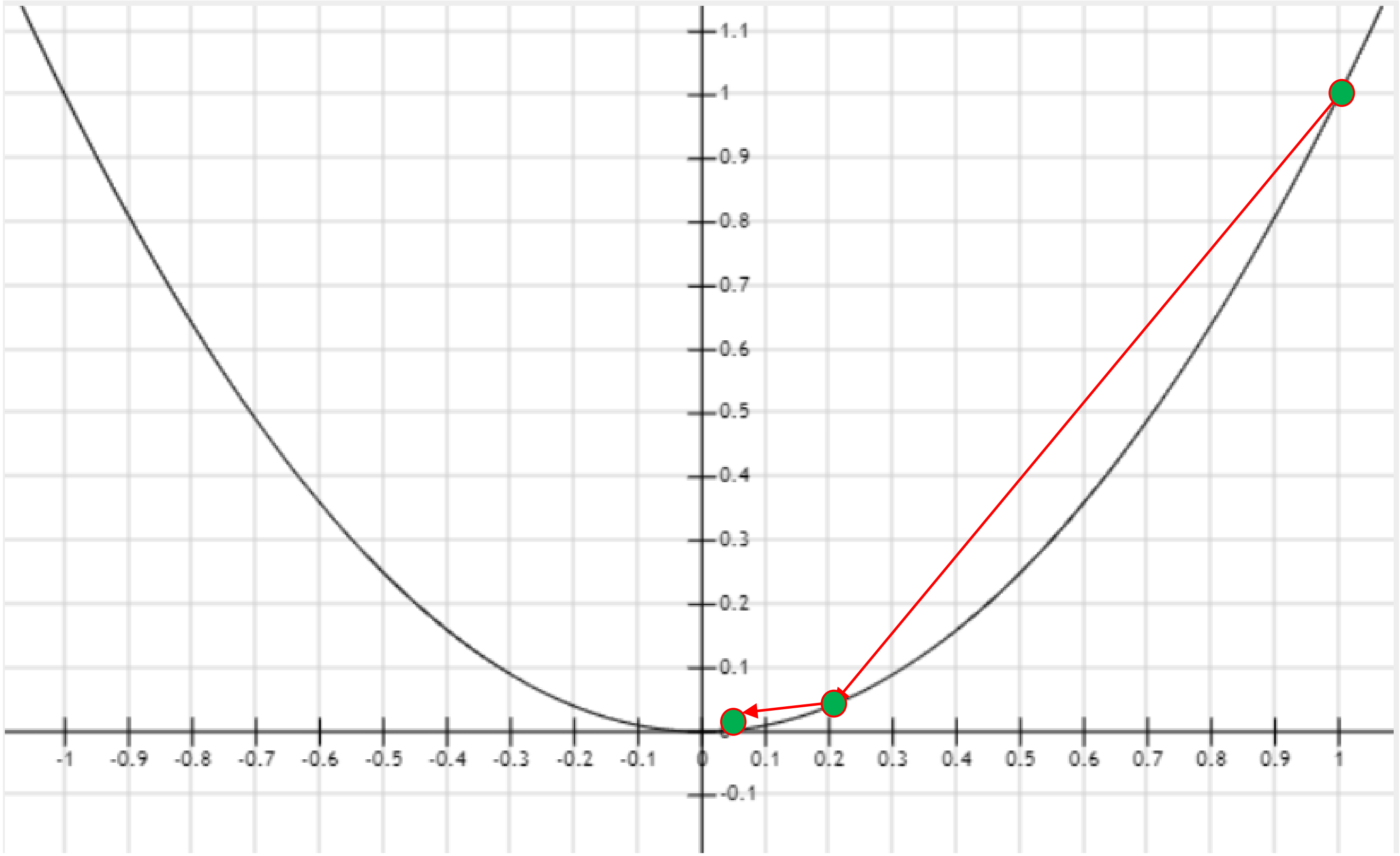
计算梯度 $\frac{\partial L(W,b)}{\partial W}$, $\frac{\partial L(W,b)}{\partial b}$;

更新参数:

$$W := W - \alpha \frac{\partial L(W,b)}{\partial W} \quad b := b - \alpha \frac{\partial L(W,b)}{\partial b}$$

3. 重复步骤2，直至L(min).

梯度下降法的细节说明



$$\mathbf{L(w) = w^2}, \boldsymbol{w_{(0)}} = \mathbf{1}, \boldsymbol{\alpha}=0.4$$

$$\boldsymbol{w := w - \alpha * \frac{\mathrm{d}L(w)}{dw}}$$

$$\boldsymbol{w_{(1)} = w_{(0)} - \alpha * \frac{dL(w)}{dw}}$$

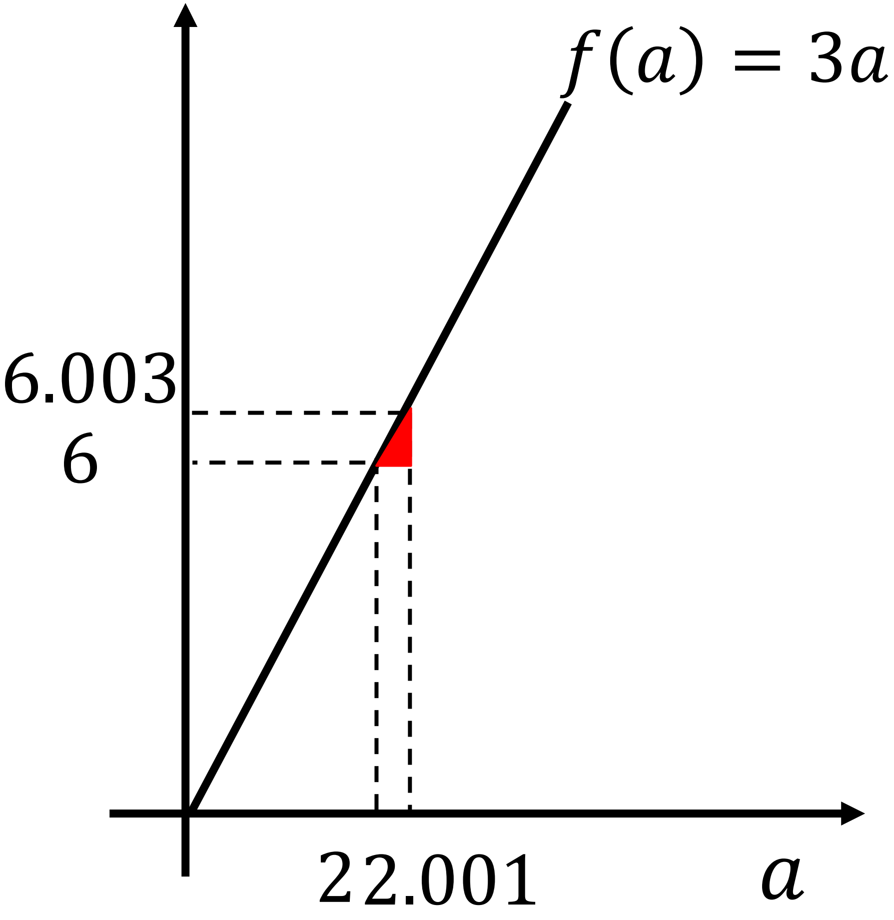
$$\mathbf{= 1 - 0.4 * 2 * 1 = 0.2}$$

$$\boldsymbol{w_{(2)}} = \mathbf{w^1} - \boldsymbol{\alpha * \frac{dL(w)}{dw}}$$

$$\mathbf{= 0.2 - 0.4 * 2 * 0.2 = 0.04}$$

导数(1)

$$\lim_{\Delta x \to 0} \frac{f(x+\Delta x)-f(x)}{\Delta x}$$



$a = 2$ $f(a) = 6$

$a = 2.001$ $f(a) = 6.003$

因此在a=2处斜率为3

$a = 5$ $f(a) = 15$

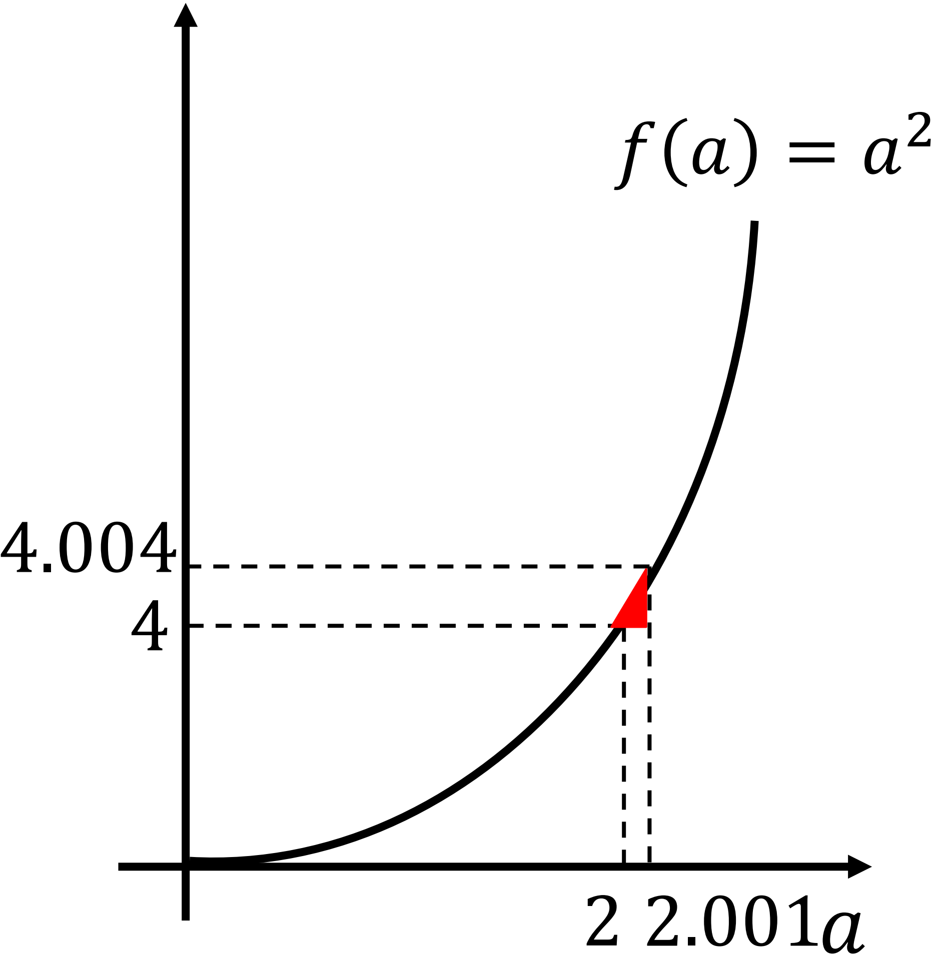
$a = 5.001$ $f(a) = 15.003$

因此在a=5处斜率为3

$$\frac{df(a)}{da} = 3 = \frac{d}{da}f(a)$$

导数(2)

$$\lim_{\Delta x \to 0} \frac{f(x+\Delta x)-f(x)}{\Delta x}$$



$$\frac{df(a)}{da}=\frac{d}{da}a^2=2a$$

$a=2$ $\quad f(a)=4$

$a=2.001$ $\quad f(a)=4.004001$

当$a=2$时，斜率为4，$\frac{df(a)}{da}=4$

$a=5$ $\quad f(a)=25$

$a=5.001$ $\quad f(a)=25.010$

(25.010001)

当$a=5$时，斜率为10，$\frac{df(a)}{da}=10$

1.4 反向传播（Backpropagation）

- Backpropagation是the backward propagation errors的缩写。
- 反向传播一般用于更新深度神经网络中的参数，具体是通过对损失函数求导来实现的。
- 反向传播包含两个环节：传播和参数更新。
- 利用计算图来解释反向传播算法。

1.4 反向传播

- 计算图导数计算

$$f(x,y,z)=(x+y)z$$

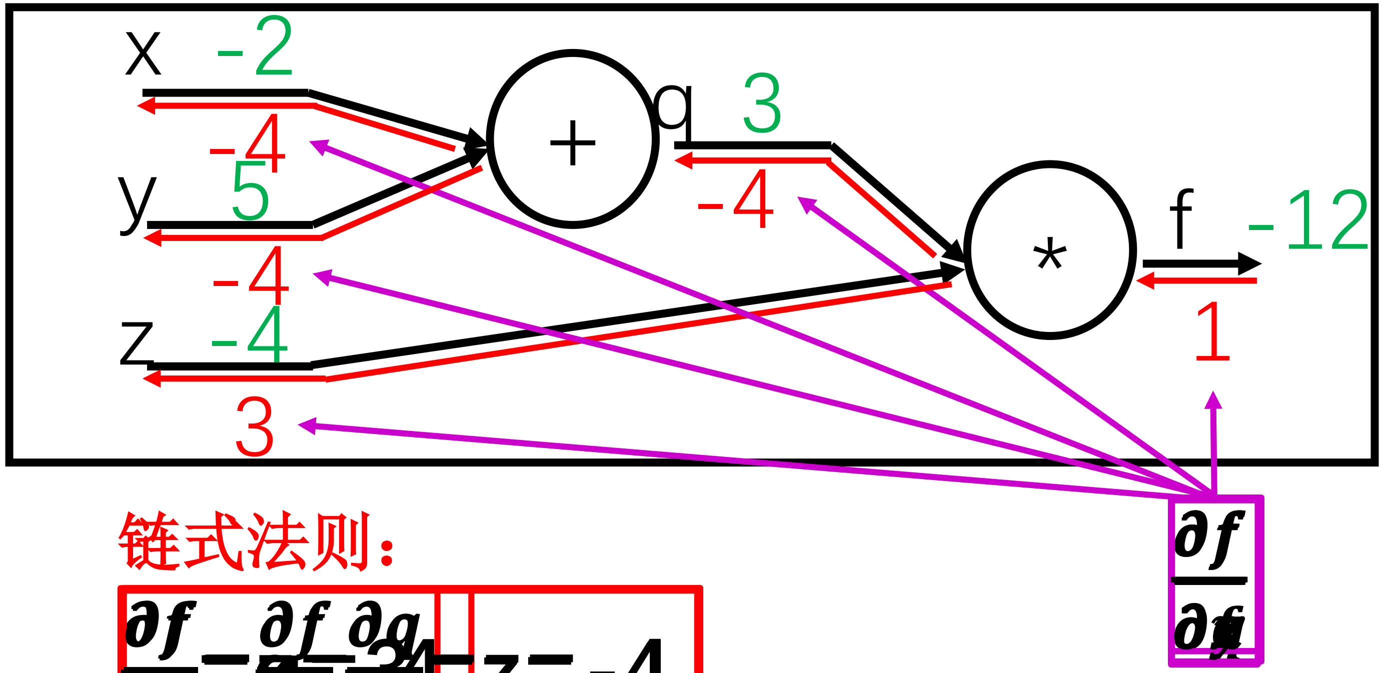
e.g. x=-2, y=5, z=-4

q= $x+y$　　$\frac{\partial q}{\partial x}$=1,　$\frac{\partial q}{\partial y}$=1

f= qz　　$\frac{\partial f}{\partial q}$=z,　$\frac{\partial f}{\partial z}$=q

Want: $\frac{\partial f}{\partial x}$　$\frac{\partial f}{\partial y}$　$\frac{\partial f}{\partial z}$

正向传播　　反向传播



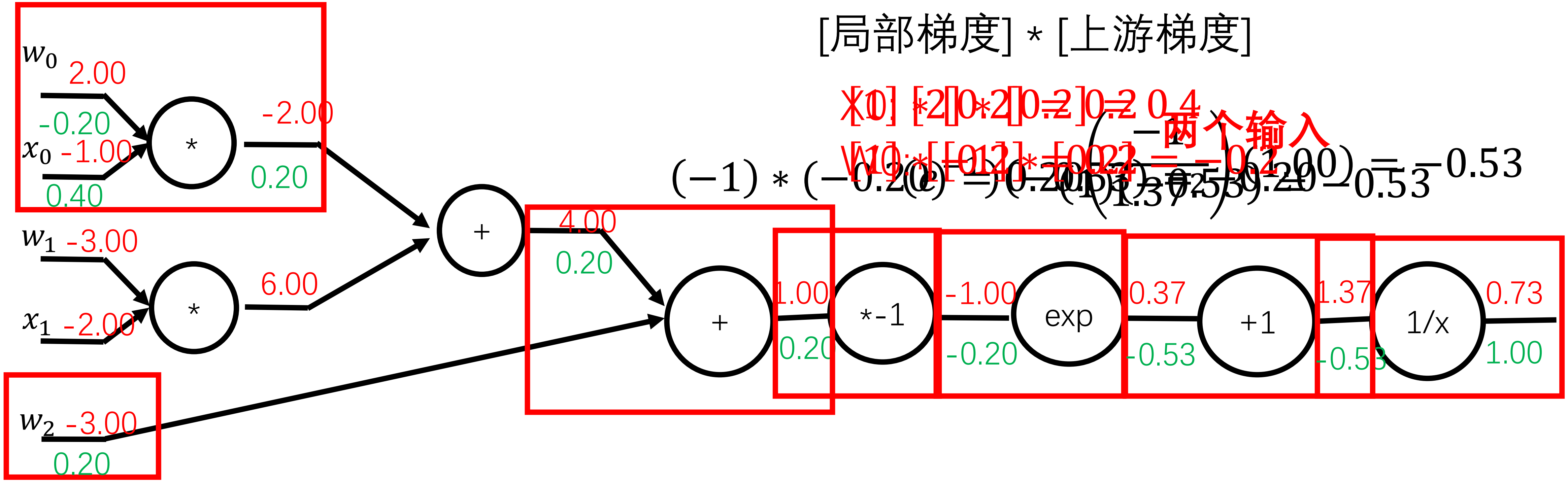
链式法则：

$\frac{\partial f}{\partial x}=\frac{\partial f}{\partial q}\frac{\partial q}{\partial x}$=z=-4

另外一个例子:

$$f(w, x) = \frac{1}{1 + e^{-(w_0 x_0 + w_1 x_1 + w_2)}}$$

[局部梯度] * [上游梯度]



$$f(x) = e^x \qquad \frac{df}{dx} = e^x$$

$$f(x) = \frac{1}{x} \qquad \frac{df}{dx} = -\frac{1}{x^2}$$

$$f(x) = ax \qquad \frac{df}{dx} = a$$

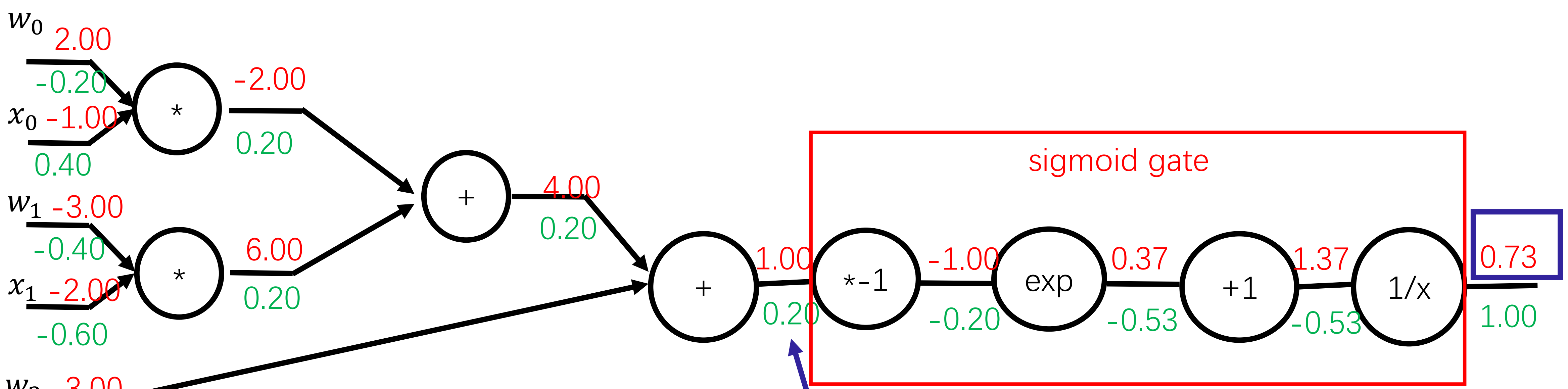
$$f_c(x) = c + x \qquad \frac{df}{dx} = 1$$

另外一个例子: f(w, x)= $\frac{1}{1+e^{-(w_0 x_0+w_1 x_1+w_2)}}$

$$f(w,x)=\frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

$$\boxed{\delta(x)=\frac{1}{1+e^{-x}}}$$ sigmoid function

$$\frac{d\delta(x)}{dx}=\frac{e^{-x}}{(1+e^{-x})^2}=\frac{1+e^{-x}-1}{1+e^{-x}}\left(\frac{1}{1+e^{-x}}\right)=(1-\delta(x))\delta(x)$$



$$\mathbf{(0.73)*(1-0.73)=0.2}$$

反向传播

- **Multiclass loss:**

$$L(W,b)=\frac{1}{N}\sum_{i=1}^{N}L_i$$

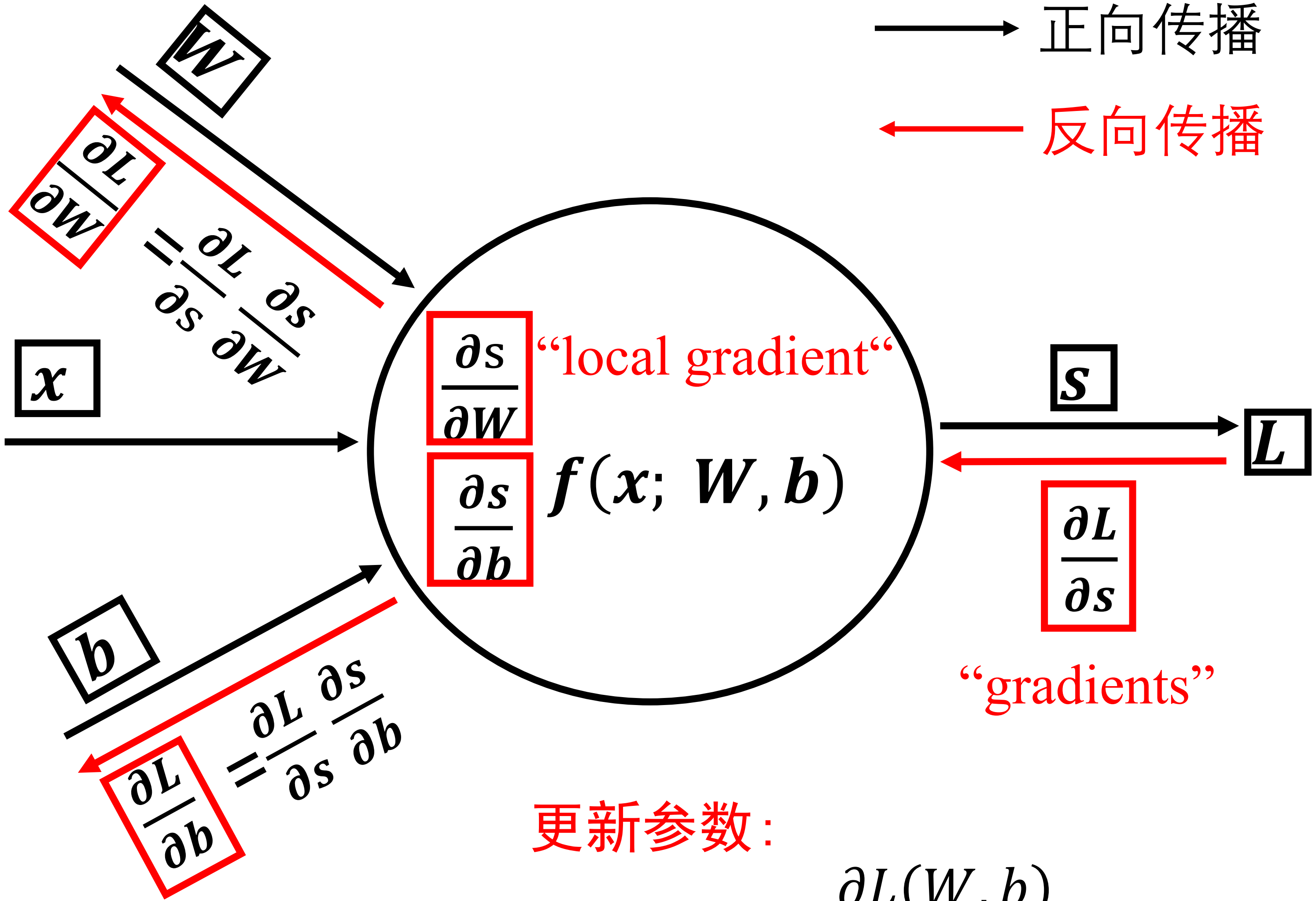
其中

Score function:

$$s=f(x_i;W,b)=Wx_i+b$$

损失函数:

$$L_i=\sum_{j\neq y_i}\max(0,s_j-s_{y_i}+1)$$



更新参数:

$$W:=W-\alpha\frac{\partial L(W,b)}{\partial W}$$

$$b:=b-\alpha\frac{\partial L(W,b)}{\partial b}$$