



۱. مقاله زیر را مطالعه بفرمایید و به سوالات زیر پاسخ دهید.

Zhuang, J.; Tang, T.; Ding, Y.; Tatikonda, S. C.; Dvornek, N.; Papademetris, X.; Duncan, J. [AdaBelief Optimizer: Adapting Stepsizes by the Belief in Observed Gradients](#). Advances in Neural Information Processing Systems 2020, 33.

ا. ایده کلی مقاله برای بهینه‌سازی را همراه با مزایای آن را توضیح دهید.

ب. تفاوت نحوه پیاده‌سازی نسبت به بهینه‌ساز Adam را به طور کامل توضیح دهید.

۲. فرض کنید در حال آموزش یک شبکه عصبی عمیق هستید. در فرآیند آموزش برای آپدیت وزن‌های شبکه از بهینه‌ساز SGD استفاده می‌کنید. در فرآیند آموزش متوجه می‌شوید مقدار تابع ضرر به کندی کاهش پیدا می‌کند و گاهی اوقات نیز افزایش می‌یابد. بعد از بررسی هایپرپارامترهای مدل متوجه می‌شوید ممکن است مقدار learning rate خیلی زیاد باشد.

ا. تحلیل شما از علت کاهش بسیار کند مقدار تابع ضرر در طول فرآیند آموزش چیست؟ به طور کامل توضیح دهید.

ب. چه راه حلی برای مشکل مطرح شده بیان می‌کنید.

ت. اگر از ابتدا مقدار learning rate را عدد کوچکی در نظر بگیریم، در آموزش مدل با چه چالش‌هایی مواجه خواهیم شد؟ همچنین چالش‌های مقدار learning rate بزرگ را نیز بیان کنید.

ث. آیا تغییری در SGD می‌توان اعمال کرد به طوری که بتوان بر مشکل به وجود آمده غلبه کرد (دو نمونه بیان کنید)؟

۳. به سایت <https://www.deeplearning.ai/ai-notes/optimization/index.html> مراجعه نمایید و به سوالات زیر پاسخ دهید:

ا. مطابق شکل برای سومین تابع، نقطه شروع را روی مقادیر  $W_1 = -1.54$ ,  $W_2 = -0.64$  تنظیم نمایید.

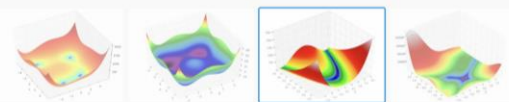
حال با مقادیر پیشفرض برای learning rate, learning rate decay که در تصویر داده شده است، آن را تا ۱۵۰ اپاک اجرا نموده و سپس نمودار حاصل را تحلیل کنید.



## تمرین سری سوم درس یادگیری عمیق

نام مدرس: دکتر محمدی  
دستیاران آموزشی مرتبط: خانم محمودی-خانم انوری  
مهلت تحویل: ۱۴۰۲/۰۱/۱۵

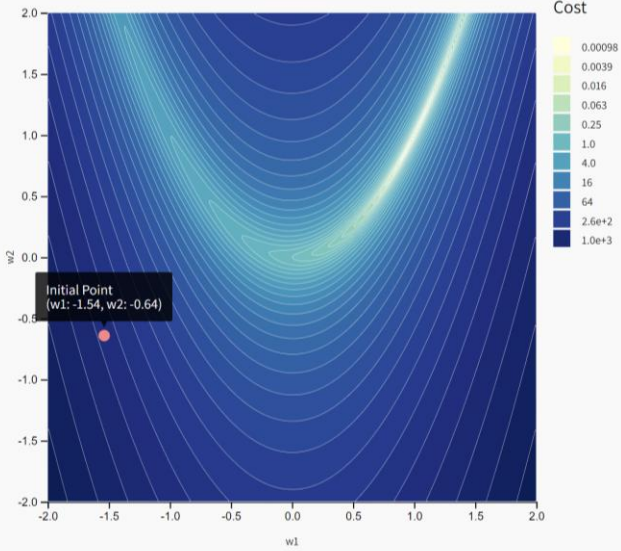
**1. Choose a cost landscape**  
Select an artificial landscape  $\mathcal{J}(w_1, w_2)$ .



**2. Choose initial parameters**  
On the cost landscape graph, drag the red dot to choose initial parameter values and thus the initial value of the cost.

**3. Choose an optimizer**  
Select the optimizer(s) and hyperparameters.

Optimizer	Learning Rate	Learning Rate Decay
<input checked="" type="checkbox"/> Gradient Descent	<input type="text" value="0.001"/>	<input type="text" value="0"/>
<input checked="" type="checkbox"/> Momentum	<input type="text" value="0.001"/>	<input type="text" value="0"/>
<input checked="" type="checkbox"/> RMSprop	<input type="text" value="0.001"/>	<input type="text" value="0"/>
<input checked="" type="checkbox"/> Adam	<input type="text" value="0.001"/>	<input type="text" value="0"/>



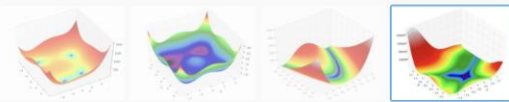
The graph below shows how the value of the cost changes through successive epochs for each optimizer.

Optimizer

ب. مطابق شکل برای چهارمین تابع، نقطه شروع را روی مقادیر  $w_1 = -0.67$ ,  $w_2 = 0.46$  تنظیم نمایید. حال مقادیر پیشفرض learning rate, learning rate decay برای ۴ الگوریتم داده شده طوری تغییر دهید که با هر کدام از بهینه‌سازها بتوان به بهترین میزان تابع ضرر دست یافت. (سعی کنید با تعداد ایپاک کمتر به مکان بهینه دست یابید).

Then, press the play button to see the optimization process (4). There's no explicit model, but you can assume that finding the cost function's minimum is equivalent to finding the best model for your task.

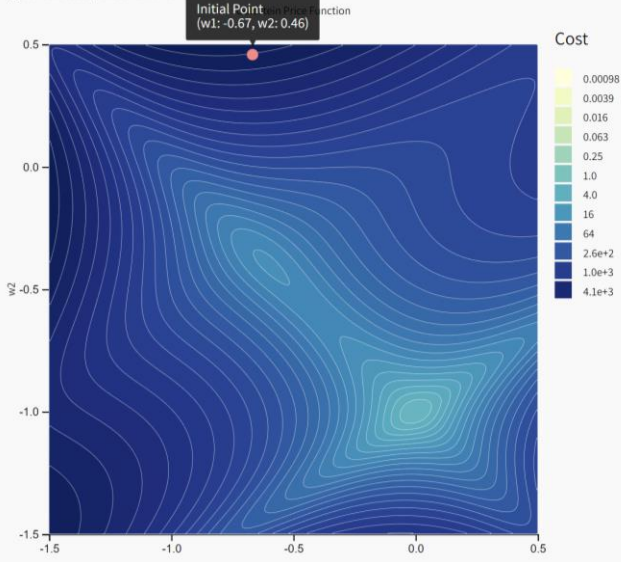
**1. Choose a cost landscape**  
Select an artificial landscape  $\mathcal{J}(w_1, w_2)$ .



**2. Choose initial parameters**  
On the cost landscape graph, drag the red dot to choose initial parameter values and thus the initial value of the cost.

**3. Choose an optimizer**  
Select the optimizer(s) and hyperparameters.

Optimizer	Learning Rate	Learning Rate Decay
<input checked="" type="checkbox"/> Gradient Descent	<input type="text" value="0.001"/>	<input type="text" value="0"/>
<input checked="" type="checkbox"/> Momentum	<input type="text" value="0.001"/>	<input type="text" value="0"/>
<input checked="" type="checkbox"/> RMSprop	<input type="text" value="0.001"/>	<input type="text" value="0"/>
<input checked="" type="checkbox"/> Adam	<input type="text" value="0.001"/>	<input type="text" value="0"/>



The graph below shows how the value of the cost changes through successive epochs for each optimizer.

Optimizer



تمرین سری سوم  
درس یادگیری عمیق

نام مدرس: دکتر محمدی  
دستیاران آموزشی مرتبط: خانم محمودی-خانم انوری  
مهلت تحویل: ۱۴۰۲/۰۱/۱۵

۴. در مسئله یادگیری مستمر، چالش اصلی از یاد رفتن کارهای پیشین است. بدین صورت که مثلاً یک مدل واحد، در کار اول با ۵ کلاس آموزش می‌بیند، سپس در کار دوم ۵ کلاس دیگر و در کار سوم ۵ کلاس جدید دیگر و.... را می‌آموزد. در هر بار آموزش، مدل تنها داده‌های مخصوص همان کار را جهت آموزش در اختیار دارد و این باعث می‌شود وزن‌هایی که برای کار پیشین آموخته بود، به سمت کار جدید تغییر پیدا کنند و همین موضوع سبب فراموشی کارهای پیشین می‌گردد. با استفاده از مفاهیم بهینه‌سازی، راهی را پیشنهاد دهید که این فراموشی به حداقل میزان ممکن برسد. راهنمایی: فرض کنید هر نرون در لایه‌های مخفی فقط یک مفهوم را یاد می‌گیرد.

۵. فرض کنید یک شبکه عصبی داریم که دارای ۲ ورودی و یک خروجی است و رابطه زیر بین ورودی و خروجی آن برقرار است ( $w_1$  و  $w_2$  و  $w_3$  و  $b$  پارامترهای قابل آموزش مدل هستند).

$$y = w_1 x_1^2 + w_2 x_2^2 + w_3 x_1 x_2 + b$$

اگر داده‌های آموزشی به صورت جدول زیر باشند و از نقطه اولیه زیر شروع کنیم، نتیجه حاصل برای  $\text{epoch}=1$  با بهینه‌ساز SGD + Momentum با  $\text{learning\_rate} = 0.1$  و  $\beta = 0.9$  به ازای  $\text{batch\_size} = 2$  محاسبه کنید. فرض کنید از تابع میانگین مربعات خطا (MSE) استفاده می‌شود. لطفاً مراحل محاسبات را به طور کامل بنویسید. برای محاسبات می‌توانید از کد numpy استفاده کنید.

$$MSE(y_i, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$w_1 = +1, w_2 = -1, w_3 = -1, b = +1$$

$X_1$	$X_2$	$y$
1	-1	10
2	0	13
0	2	11
-1	1	4



تمرین سری سوم  
درس یادگیری عمیق

نام مدرس: دکتر محمدی  
دستیاران آموزشی مرتبط: خانم محمودی-خانم انوری  
مهلت تحویل: ۱۴۰۲/۰۱/۱۵

۶. warmup یکی از تکنیک‌هایی است که اخیراً در بهینه‌سازی شبکه‌های عصبی استفاده می‌گردد.

ا. این تکنیک و مزایای حاصل از آن را توضیح دهید.

ب. این روش با learning rate decay چه تفاوتی دارد؟

ت. کد بخش [warm up](#) کتاب را اجرا نموده و تحلیل خود را از مشاهدات بدست آمده بنویسید.

لطفاً سند قوانین انجام و تحویل تمرین‌های درس را مطالعه و موارد خواسته شده را رعایت فرمایید

موفق و سلامت باشید

