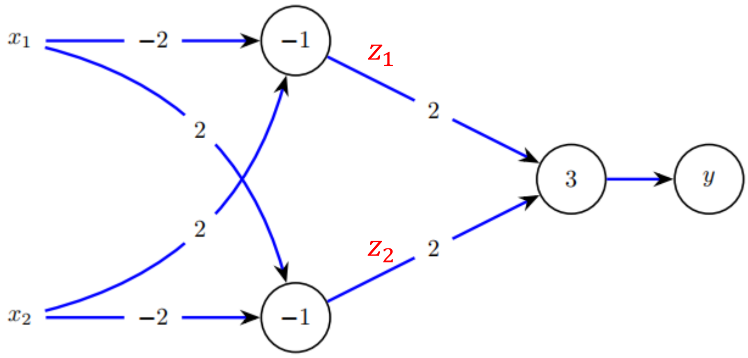
۱ شبکه‌های عصبی



.۱

اگر $\alpha = 2x_2 - 2x_1$ را در نظر بگیریم، بر حسب مقادیر مختلف α می‌توان خروجی را مشخص کرد (منظور از θ تابع فعالسازی آستانه می‌باشد):

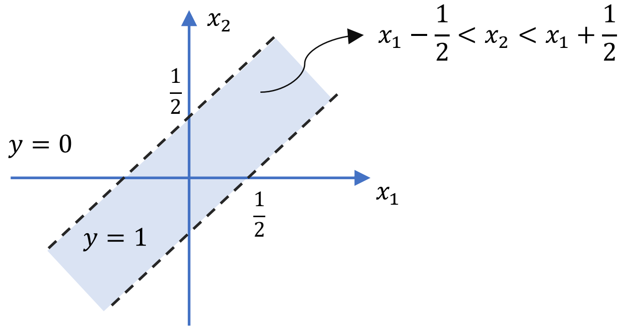
- $\alpha \geq 1 \Rightarrow z_1 = \theta\big(\alpha - (-1)\big) = 1$
  $\alpha \geq 1 \Rightarrow 0 \geq 1 - \alpha \Rightarrow z_2 = \theta\big(-\alpha - (-1)\big) = 0$
  $\Rightarrow y = \theta(2z_1 + 2z_2 - 3) = 0$

- $1 > \alpha > -1 \Rightarrow z_1 = \theta\big(\alpha - (-1)\big) = 1$
  $1 > \alpha > -1 \Rightarrow 1 - \alpha > 0 \Rightarrow z_2 = \theta\big(-\alpha - (-1)\big) = 1$
  $\Rightarrow y = \theta(2z_1 + 2z_2 + 3) = 1$

- $-1 \geq \alpha \Rightarrow z_1 = \theta\big(\alpha - (-1)\big) = 0$
  $-1 \geq \alpha \Rightarrow z_2 = \theta\big(-\alpha - (-1)\big) = 1$
  $\Rightarrow y = \theta(2z_1 + 2z_2 - 3) = 0$

بنابراین بازه‌ای که در آن خروجی (y) مقدار یک خواهد داشت، به صورت $1 > \alpha > -1$ مشخص می‌شود:

$$-1 < \alpha < 1 \Rightarrow -1 < 2x_2 - 2x_1 < +1 \Rightarrow x_1 - \frac{1}{2} < x_2 < x_1 + \frac{1}{2}$$



.۲

جدول مقداردهی باینری به متغیرها به صورت زیر است:

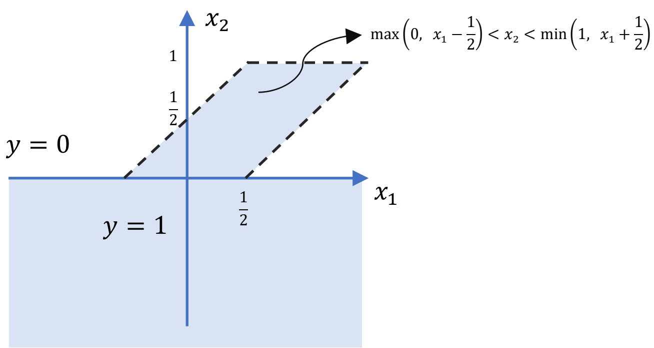
x_1	x_2	$\alpha = 2x_2 - 2x_1$	$z_1 = \theta(\alpha + 1)$	$z_1 = \theta(-\alpha + 1)$	$y = \theta(2z_1 + 2z_2 - 3)$
0	0	0	1	1	1
0	1	2	1	0	0
1	0	-2	0	1	0
1	1	0	1	1	1

با توجه به جدول فوق، به نظر می‌رسد تابع منطقی پیاده‌سازی شده به صورت $\overline{(x_1\ xor\ \ x_2)}$ می‌باشد.

.۳

در قسمت قبل دیدیم که وقتی $x_1 - \frac{1}{2} < x_2 < x_1 + \frac{1}{2}$ خروجی N برابر یک خواهد شد. در شبکه جدید، مقدار متغیر x_2 با خروجی شبکه سنجیده می‌شود. بنابراین مقدار خروجی را در چند حالت بررسی می‌کنیم (توجه کنید که در هر شرایطی $0 \le N \le 1$ می‌باشد):

- $x_2 < 0 \Rightarrow -x_2 + N > 0 \Rightarrow z = 1$
- $x_2 > 1 \Rightarrow -x_2 + N < 0 \Rightarrow z = 0$
- $\max\left(0,\ \ x_1 - \frac{1}{2}\right) < x_2 < \min\left(1,\ \ x_1 + \frac{1}{2}\right) \Rightarrow N = 1 \Rightarrow -x_2 + N = -x_2 + 1 > 0 \Rightarrow z = 1$



۲ مدل IF تعمیم‌یافته

.۱

$$\frac{dV}{dt} = a(I - I_0) + b(V - V_0)^2 = 0 \Rightarrow (V - V_0)^2 = \frac{a}{b}(I_0 - I) \Rightarrow V = V_0 \pm \sqrt{\frac{a}{b}(I_0 - I)}$$

تا زمانی که $I_0 > I$ سیستم دو نقطه ثابت دارد. اگر جریان ورودی تغییر یابد به طوری که $I_0 = I$، در این صورت یک نقطه ثابت وجود خواهد داشت و اگر $I_0 < I$ شود، هر دو نقطه ثابت محو خواهند شد.

.۲

برای بررسی پایداری، کافیست علامت مشتق عبارت $a(I - I_0) + b(V - V_0)^2$ را در اطراف نقاط ثابت بررسی کنیم:

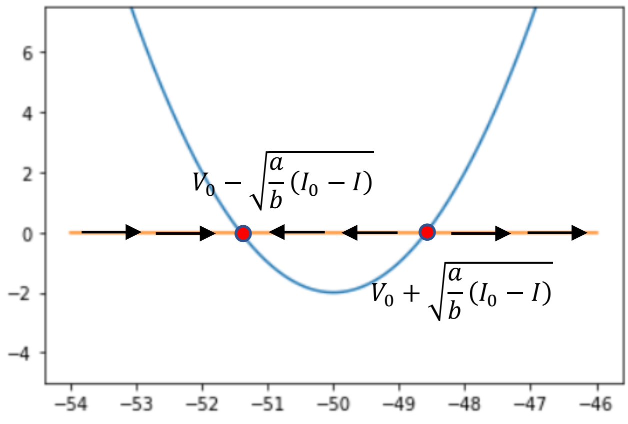
$$\frac{d}{dV}[a(I - I_0) + b(V - V_0)^2] = 2b(V - V_0)$$

$$V = V_0 + \sqrt{\frac{a}{b}(I_0 - I)} \Rightarrow 2b(V - V_0) = 2b\sqrt{\frac{a}{b}(I_0 - I)} > 0 \Rightarrow unstable\ fixed\ point$$

$$V = V_0 - \sqrt{\frac{a}{b}(I_0 - I)} \Rightarrow 2b(V - V_0) = -2b\sqrt{\frac{a}{b}(I_0 - I)} < 0 \Rightarrow stable\ fixed\ point$$

.۳

مدل ارائه شده می‌تواند تا حد خیلی خوبی مشابه یک نورون واقعی عمل کند. اولاً نقطه $V_0 - \sqrt{\frac{a}{b}(I_0 - I)}$ را می‌توان ولتاژ استراحت نورون در نظر گرفت چرا که انحرافات کوچک از این نقطه باعث می‌شود تا دینامیک معادلات به صورت خودکار ولتـــاژ را به این نقطه بازگرداند. از طرفی $V_0 + \sqrt{\frac{a}{b}(I_0 - I)}$ را می‌توان شبیه به ولتاژ آستانه دانست. زیرا تا زمانی که ولتاژ کمتر از این مقدار است، به صورت خودکار به سمت نقطه ثابت پایدار جذب می‌شود اما همین که از این مقدار آستانه بیشتر شد، به علت عدم تعادل نقطه ثابت، ولتاژ به سمت مثبت شدن سرعت زیادی می‌گیرد و دیگر به این نقطه باز نمی‌گردد. البته این مدل هنوز به طور کامل رفتار یک نورون را نشان نمی‌دهد چرا که بعد از اسپایک، مکانیزم مناسبی برای کاهش دوباره ولتاژ به مقدار استراحت ندارد. برای این منظور می‌توان یک ولتاژ reset تعریف نمود یا از متغیر کمکی دومی و تحلیل صفحه فاز استفاده کرد.



۴-

$$\frac{dV}{dt} = F(V,u) = -V + V^2 - u = 0 \Rightarrow u = V^2 - V$$

$$\frac{du}{dt} = G(V,u) = V - u = 0 \Rightarrow u = V$$

$$\begin{cases} u = V^2 - V \\ u = V \end{cases} \Rightarrow V = V^2 - V \Rightarrow (V_1, u_1) = (0{,}0), (V_2, u_2) = (2{,}2)$$

۵.

$$\frac{\partial F(V,u)}{\partial V} = -1 + 2V, \qquad \frac{\partial F(V,u)}{\partial u} = -1$$

$$\frac{\partial G(V,u)}{\partial V} = 1, \qquad \frac{\partial G(V,u)}{\partial u} = -1$$

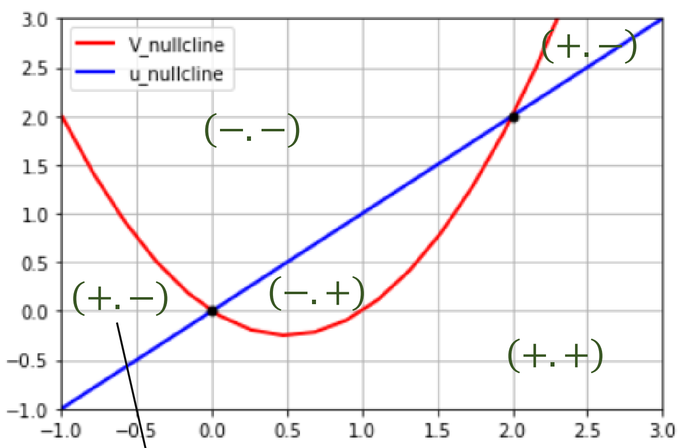
$$(V_1, u_1) = (0{,}0) \Rightarrow A = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} \Rightarrow (-1-\lambda)(-1-\lambda) + 1 = 0 \Rightarrow \lambda^2 + 2\lambda + 2 = 0$$
$$\Rightarrow \lambda = -1 \pm j$$

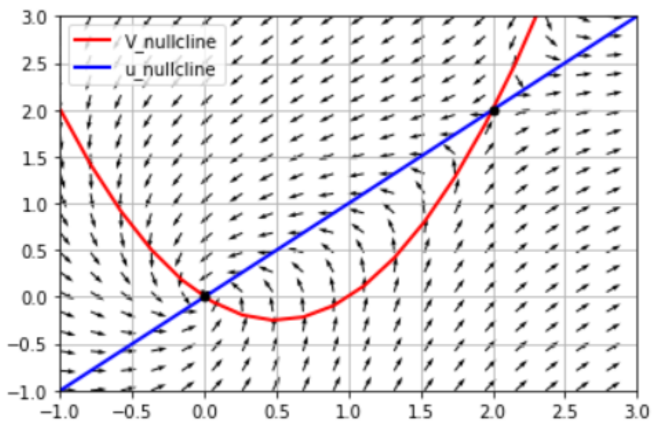
با توجه به مثبت نبودن قسمت حقیقی دو مقدار ویژه، این نقطه تعادل پایدار است.

$$(V_2, u_2) = (2{,}2) \Rightarrow A = \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} \Rightarrow (3-\lambda)(-1-\lambda) + 1 = 0 \Rightarrow \lambda^2 - 2\lambda - 2 = 0 \Rightarrow \lambda = \frac{2 \pm \sqrt{12}}{2}$$
$$= 1 \pm \sqrt{3}$$

با توجه به مثبت شدن مقدار یکی از این مقادیر ویژه، این نقطه تعادل ناپایدار (زینی) است.

۶.





$$\boxed{-V + V^2 > u > V \Rightarrow \frac{dV}{dt} = -V + V^2 - u > 0\,, \qquad \frac{du}{dt} = V - u < 0}$$

توجه: قبل از پاسخ به سوال بعد، توجه کنید که این مدل هنوز مدل کاملی نیست چرا که مکانیزمی برای reset و بازگرداندن ولتاژ به مقدار استراحت ندارد. بنابراین برای تحلیل این مدل در مسئله بعدی یک مکانیزم ریست فرضی به صورت زیر تعریف می‌کنیم:

if $V > 6$ then:

$V = -1$

$u = u + 1$

.۷

مشابه قسمت قبل، مجدداً معادلات را نوشته و نقاط تعادل را می‌یابیم:

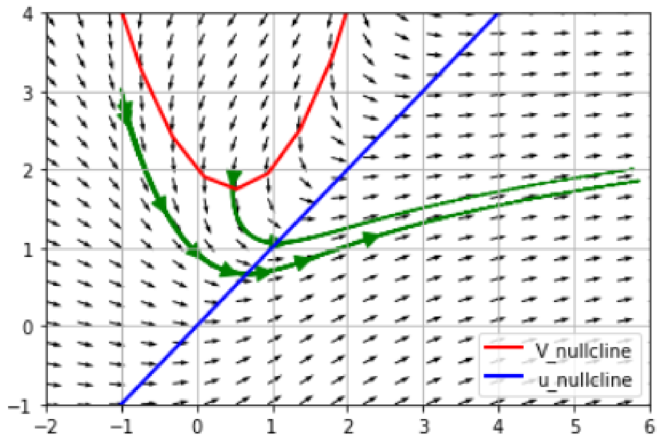
$$\begin{cases} u = V^2 - V + I_0 \\ u = V \end{cases} \Rightarrow V = V^2 - V + I_0 \Rightarrow V^2 - 2V + I_0 = 0$$

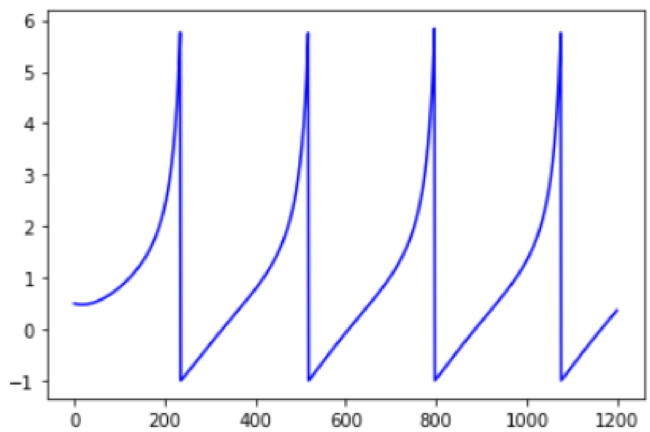
برای آن که معادله درجه ۲ فوق جواب داشته باشد، باید داشته باشیم:

$$4 - 4I_0 \geq 0 \Rightarrow I_0 \leq 1$$

بنابراین اگر جریان ورودی از 1 بیشتر شود، این مدل نقطه تعادلی ندارد.

اگر مکانیزم resetای که در بالا معرفی کردیم را استفاده نکنیم، مدل ناپایدار است و نمی‌تواند مدل کننده یک نورون واقعی باشد. اما با در نظر گرفتن این مکانیزم، معادلات ما رفتار مناسبی خواهند داشت. به عبارت دیگر، اگر جریان از 1 بیشتر باشد، مدل دچار bifurcation شده و هیچ نقطه تعادلی وجود نخواهد داشت. عدم وجود نقطه تعادل به همراه در نظر گرفتن مکانیزم reset، با رفتار نورون واقعی همخوانی دارد چرا که در صورت تزریق جریان ثابت زیاد، باعث اسپایک زدن متناوب نورون خواهد شد. نتایج شبیه سازی نیز این مسئله را تایید می‌کنند (منحنی‌های سبز رنگ trajectory را نشان می‌دهند):





۳ محاسبه‌ی پاسخ نورون به محرک

.۱

معادله دیفرانسیلی که با آن سر و کار خواهیم داشت، به صورت زیر است:

$$\tau_m \frac{dv}{dt} = -(v - v_0) + RI(t) = -(v - v_0) + R\left[\frac{q}{\tau_s} e^{-\frac{t-t_f}{\tau_s}} u\left(t - t_f\right)\right]$$

با توجه به تکنیک اضافه کردن عامل انتگرال ساز و همچنین مباحث مطرح شده در این لینک، می‌توان پاسخ معادله فوق را به صورت زیر ارائه کرد:

$$p(t) = \frac{1}{\tau_m},\ \ g(t) = \frac{1}{\tau_m}[v_0 + RI(t)] \Rightarrow \mu(t) = e^{\int p(t)dt}$$

$$\Rightarrow v(t) = \frac{c + \frac{1}{\tau_m}\int_{t_f}^{t} e^{\frac{s-t_f}{\tau_m}}(RI(s) + v_0)ds}{e^{\frac{t-t_f}{\tau_m}}} = e^{-\frac{t-t_f}{\tau_m}}\left[c + \frac{1}{\tau_m}\int_{t_f}^{t} e^{\frac{s-t_f}{\tau_m}}(RI(s) + v_0)ds\right]$$

$$= e^{-\frac{t-t_f}{\tau_m}}\left[c + \frac{1}{\tau_m}\int_{t_f}^{t} e^{\frac{s-t_f}{\tau_m}}\left(\frac{Rq}{\tau_s} e^{-\frac{s-t_f}{\tau_s}} u\left(s - t_f\right) + v_0\right)ds\right] =$$

$$= e^{-\frac{t-t_f}{\tau_m}}\left[c + \frac{1}{\tau_m}\int_{t_f}^{t} \frac{Rq}{\tau_s} e^{\frac{s-t_f}{\tau_m} - \frac{s-t_f}{\tau_s}} + v_0 e^{\frac{s-t_f}{\tau_m}} ds\right]$$

$$= e^{-\frac{t-t_f}{\tau_m}}\left[c + \frac{1}{\tau_m}\left[\frac{Rq}{\tau_s}\frac{\tau_m\tau_s}{\tau_s - \tau_m}\ (e^{\frac{t(\tau_s-\tau_m)-t_f(\tau_s-\tau_m)}{\tau_m\tau_s}} - 1) + v_0\tau_m\left(e^{\frac{t-t_f}{\tau_m}} - 1\right)\right]\right]$$

$$\Rightarrow v(t) = ce^{-\frac{t-t_f}{\tau_m}} + \frac{Rq}{\tau_s - \tau_m}\left(e^{-\frac{t-t_f}{\tau_s}} - e^{-\frac{t-t_f}{\tau_m}}\right) + v_0\left(1 - e^{-\frac{t-t_f}{\tau_m}}\right)$$

پس ولتاژ غشای نورون در زمان t را می‌توان طبق رابطه بالا به دست آورد. c یک مقدار ثابت است که با توجه به شرایط اولیه تعیین می‌شود. با توجه به این که ولتاژ ابتدایی نورون قبل از رسیدن اسپایک در حالت استراحت v_0، قرار دارد، می‌توان مقدار c را به دست آورد:

$$v_0 = v\left(t_f\right) = ce^{-\frac{t_f-t_f}{\tau_m}} + \frac{Rq}{\tau_s - \tau_m}\left(e^{-\frac{t_f-t_f}{\tau_s}} - e^{-\frac{t_f-t_f}{\tau_m}}\right) + v_0\left(1 - e^{-\frac{t_f-t_f}{\tau_m}}\right) = c \Rightarrow c = v_0$$

$$\Rightarrow v(t) = \frac{Rq}{\tau_s - \tau_m}\left(e^{-\frac{t-t_f}{\tau_s}} - e^{-\frac{t-t_f}{\tau_m}}\right) + v_0$$

.۲

$$\tau_s = \tau_m \Rightarrow v(t) = \lim_{\tau_s \to \tau_m} \frac{Rq}{\tau_s - \tau_m}\left(e^{-\frac{t-t_f}{\tau_s}} - e^{-\frac{t-t_f}{\tau_m}}\right) + v_0 = \frac{Rq}{\tau_m^2}\left(t - t_f\right)e^{-\frac{t-t_f}{\tau_m}} + v_0$$

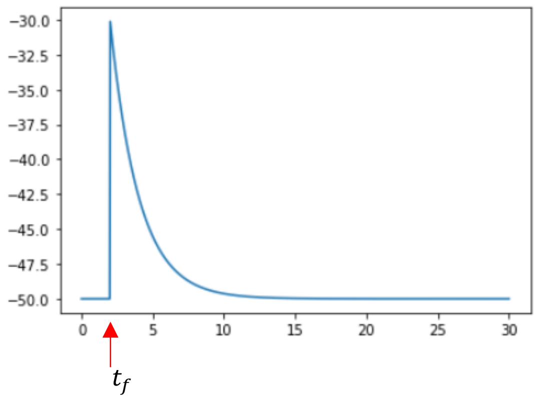
.۳

$$\tau_s \to 0 \Rightarrow v(t) = \lim_{\tau_s \to 0} \frac{Rq}{\tau_s - \tau_m}\left(e^{-\frac{t-t_f}{\tau_s}} - e^{-\frac{t-t_f}{\tau_m}}\right) + v_0 = \frac{Rq}{-\tau_m}\left(0 - e^{-\frac{t-t_f}{\tau_m}}\right) + v_0 = \frac{Rq}{\tau_m} e^{-\frac{t-t_f}{\tau_m}} + v_0$$

.۴

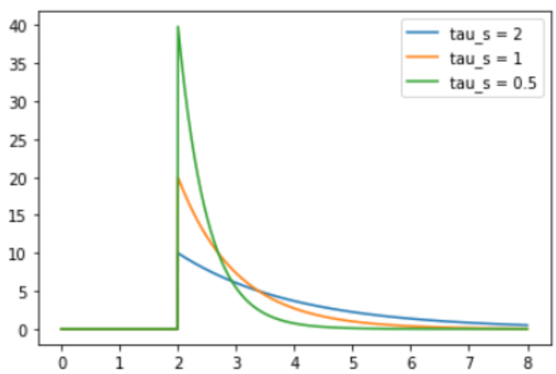
برای پاسخ به این بخش، لازم است تا بار دیگر معادله به دست آمده برای $v(t)$ را در شرایط $\tau_s = 0$ بازنگری کنیم:

$$v(t) = \begin{cases} v_0, & t < t_f \\ \frac{Rq}{\tau_m} e^{-\frac{t-t_f}{\tau_m}} + v_0, & t \geq t_f \end{cases}$$



همانطور که در شکل فوق نیز نشان داده شده است، ولتاژ نورون تا قبل از رسیدن جریان ورودی، روی مقدار استراحت v_0 قرار داشته است. با رسیدن جریان ورودی و تزریق بار q به فضای داخل نورون، ولتاژ غشاء نورون پرش ناگهانی داشته و به میزان $\frac{Rq}{\tau_m}$ زیاد می‌شود. به عبارت دقیق‌تر اگر به ولتاژ در لحظه $t = t_f$ نگاه کنیم، عبارت $v(t_f) = \frac{Rq}{\tau_m} + v_0$ ظاهر می‌شود که به اندازه $\frac{Rq}{\tau_m}$ از سطح ولتاژ استراحت بالاتر است. در زمان‌های بعد از t_f ولتاژ نورون به صورت نمایی روند نزولی می‌گیرد تا دوباره به شرایط استراحت ($v(t) = v_0$) بازگردد.

این پاسخ ارتباط بسیار نزدیکی با پاسخ به جریان ضربه دارد. برای درک این موضوع، لازم است تا شکل جریان ورودی و وابستگی آن به τ_s را در نظر بگیریم: $I(t) = \frac{q}{\tau_s} e^{-\frac{t-t_f}{\tau_s}} u(t - t_f)$. پارامتر τ_s هم بزرگی مقدار جریان را تعیین می‌کند و هم عرض پالس را نشان می‌دهد به طوری که هرچه τ_s به صفر نزدیک تر شود، عرض پالس کمتر شده و مقدار دامنه آن بیشتر می‌شود. بنابراین در حالت حدی، یک جریان ورودی با دامنه خیلی زیاد و عرض خیلی کوچک خواهیم داشت. در شکل زیر، جریان ورودی برای چند مقدار τ_s رسم شده است. می‌بینیم که با کم شدن τ_s، جریان ورودی به تابع دلتا نزدیک‌تر شده و در نتیجه رفتار ولتاژ به پاسخ ضربه نزدیک‌تر می‌شود.



۴ قوانین یادگیری

.۱

برای یافتن نقطه ثابت، لازم است تا به مشتق v_i نگاه کنیم:

$$v_i = g(\sum_{j=1}^{N} w_{ij}v_j) \Rightarrow \frac{dv_i}{dt} = g'(\sum_{j=1}^{N} w_{ij}v_j) \times \frac{d}{dt}\sum_{j=1}^{N} w_{ij}v_j = g'(\sum_{j=1}^{N} w_{ij}v_j)\sum_{j=1}^{N} v_j \frac{dw_{ij}}{dt}$$

$$= g'(\sum_{j=1}^{N} w_{ij}v_j)\sum_{j=1}^{N} v_j\big[\gamma(v_i - v_\theta)v_j\big] = g'(\sum_{j=1}^{N} w_{ij}v_j)\gamma(v_i - v_\theta)\sum_{j=1}^{N} v_j^2$$

$$\Rightarrow \frac{dv_i}{dt} = \gamma\, g'(\sum_{j=1}^{N} w_{ij}v_j)(\sum_{j=1}^{N} v_j^2)(v_i - v_\theta)$$

با توجه به معادله فوق برای مشتق v_i می‌بینیم که v_θ نقطه ثابت این معادله است. چرا که با قرار دادن $v_i = v_\theta$ سمت راست معادله صفر شده، و در نتیجه $\frac{dv_i}{dt} = 0$ خواهد بود. لذا v_i روی مقدار v_θ ثابت خواهد ماند.

.۲

به طور کلی برای بررسی وضعیت پایداری نقطه تعادل معادله دیفرانسیلی به فرم $\frac{dv}{dt} = a(v-b)$، لازم است به جهت شیب عبور از صفر در نقطه تعادل بنگریم. به طور دقیق‌تر اگر $a > 0$ باشد، این نقطه تعادل ناپایدار و اگر $a < 0$ باشد، نقطه تعادل پایدار است.

با این تفاسیر، برای بررسی پایداری، لازم است علامت $\gamma\, g'(\sum_{j=1}^{N} w_{ij}v_j)(\sum_{j=1}^{N} v_j^2)$ را مورد بررسی قرار دهیم. عبارت $\sum_{j=1}^{N} v_j^2$ که همیشه مثبت است. همچنین تابع g معمولاً تابعی به صورت ReLU یا sigmoid تعریف می‌شود که این توابع همواره صعودی هستند و مشتق آن‌ها مثبت است. بنابراین در معادله فوق، γ تعیین کننده پایداری است به طوری که اگر $\gamma > 0$ باشد، نقطه ثابت ناپایدار است و در غیر اینصورت نقطه ثابت پایدار خواهد بود.

---

۵ یادگیری بهینه و ارزش حالت‌های میانی

.۱

$$V(D) = 0.5 \times r_1 + 0.5 \times V(E) = \frac{V(E)}{2} + 1$$

$$V(E) = 0.7 \times r_2 + 0.3 \times V(D) = \frac{3\,V(D)}{10} + 5.6$$

$$\Rightarrow V(D) = \frac{1}{2}\left(\frac{3\,V(D)}{10} + 5.6\right) + 1 \Rightarrow \frac{17}{20}V(D) = 3.8 \Rightarrow V(D) = \frac{76}{17} = 4.47\ ,\ \ V(E) = 6.94$$

$$V(B) = 0.3V(C) + 0.7V(D) \Rightarrow V(B) - 0.3V(C) = 0.7V(D)$$

$$V(C) = 0.4V(B) + 0.6V(E) \Rightarrow -0.4V(B) + V(C) = 0.6V(E)$$

$$\Rightarrow \begin{bmatrix} 1 & -0.3 \\ -0.4 & 1 \end{bmatrix}\begin{bmatrix} V(B) \\ V(C) \end{bmatrix} = \begin{bmatrix} 0.7V(D) \\ 0.6V(E) \end{bmatrix} \Rightarrow \begin{bmatrix} V(B) \\ V(C) \end{bmatrix} = \frac{1}{0.88}\begin{bmatrix} 1 & 0.3 \\ 0.4 & 1 \end{bmatrix}\begin{bmatrix} 0.7V(D) \\ 0.6V(E) \end{bmatrix} = \frac{1}{0.88}\begin{bmatrix} 1 & 0.3 \\ 0.4 & 1 \end{bmatrix}\begin{bmatrix} 3.129 \\ 4.164 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} V(B) \\ V(C) \end{bmatrix} = \begin{bmatrix} 4.975 \\ 6.154 \end{bmatrix}$$

$$V(A) = 0.4 \times V(B) + 0.6 \times V(C) = 1.99 + 3.69 = 5.68$$

۲.

برای پاسخ به این سوال، لازم است تا در هر state پاداش مورد انتظار از هر اکشن را با پاداش میانگین مقایسه کنیم:

در حالت A، با پاداش متوسط $V(A) = 5.68$:

- پاداش مورد انتظار در صورت انتخاب مسیر چپ : $V(B) = 4.975$ (کمتر از میانگین)
- پاداش مورد انتظار در صورت انتخاب مسیر راست : $V(C) = 6.154$ (بیشتر از میانگین)
- بنابراین در مورد حالت A، احتمال اکشن راست تقویت شده و اکشن چپ تضعیف می‌شود.

در حالت B، با پاداش متوسط $V(B) = 4.975$:

- پاداش مورد انتظار در صورت انتخاب مسیر چپ : $V(D) = 4.47$ (کمتر از میانگین)
- پاداش مورد انتظار در صورت انتخاب مسیر راست : $V(C) = 6.154$ (بیشتر از میانگین)
- بنابراین در مورد حالت B، احتمال اکشن راست تقویت شده و اکشن چپ تضعیف می‌شود.

در حالت C، با پاداش متوسط $V(C) = 6.154$:

- پاداش مورد انتظار در صورت انتخاب مسیر چپ : $V(B) = 4.975$ (کمتر از میانگین)
- پاداش مورد انتظار در صورت انتخاب مسیر راست : $V(E) = 6.94$ (بیشتر از میانگین)
- بنابراین در مورد حالت C، احتمال اکشن راست تقویت شده و اکشن چپ تضعیف می‌شود.

در حالت D، با پاداش متوسط $V(D) = 4.47$:

- پاداش مورد انتظار در صورت انتخاب مسیر چپ : $V(r_1) = 2$ (کمتر از میانگین)
- پاداش مورد انتظار در صورت انتخاب مسیر راست : $V(E) = 6.94$ (بیشتر از میانگین)
- بنابراین در مورد حالت E، احتمال اکشن راست تقویت شده و اکشن چپ تضعیف می‌شود.

در حالت E، با پاداش متوسط $V(E) = 6.94$:

- پاداش مورد انتظار در صورت انتخاب مسیر چپ : $V(D) = 4.47$ (کمتر از میانگین)
- پاداش مورد انتظار در صورت انتخاب مسیر راست : $V(r_2) = 8$ (بیشتر از میانگین)
- بنابراین در مورد حالت E، احتمال اکشن راست تقویت شده و اکشن چپ تضعیف می‌شود.

بنابراین در تمامی حالات، احتمال رفتن به سمت راست افزایش یافته و احتمال رفتن به چپ کاهش می‌یابد. البته این مسئله از قبل نیز قابل پیش‌بینی بود چرا که پاداش انتهایی شاخه سمت راست (r_2) بیشتر از سمت چپ (r_1) می‌باشد.

۶ یادگیری تقویتی و استراتژی بهینه

.۱

برای پاسخ به این سوال، ابتدا امید ریاضی پاداش در A را می‌نویسیم. برای این منظور به امید پاداش در حالات میانی نیز احتیاج داریم. لذا از پایین درخت شروع کرده و به سمت ریشه حرکت می‌کنیم:

$$V(B) = \mathbb{E}[r] = 2p \times r_3 + (1-2p) \times r_4 = \frac{6p}{1+p^2} + \frac{4(1-2p)}{1+p^2} = \frac{4-2p}{1+p^2}$$

$$V(C) = 3p \times r_2 + (1-3p) \times r_1 = \frac{6p}{1+p^2} + \frac{(1-3p)}{1+p^2} = \frac{1+3p}{1+p^2}$$

$$V(A) = p \times V(B) + (1-p) \times V(C) = \frac{4p-2p^2}{1+p^2} + \frac{(1-p)(1+3p)}{1+p^2} = \frac{-5p^2+6p+1}{1+p^2}$$

حال برای آن که امید پاداش در A حداکثر شود، نسبت به p مشتق می‌گیریم:

$$\frac{dV(A)}{dp} = 0 \Rightarrow \frac{d}{dp}\frac{-5p^2+6p+1}{1+p^2} = \frac{(6-10p)(1+p^2)-(2p)(-5p^2+6p+1)}{(1+p^2)^2}$$
$$= \frac{6-10p+6p^2-10p^3+10p^3-12p^2-2p}{(1+p^2)^2} = \frac{6-12p-6p^2}{(1+p^2)^2} = 0 \Rightarrow p = -1 \pm \sqrt{2}$$

که تنها جواب $p = -1+\sqrt{2} \approx 0.4$ مورد قبول است.

.۲

مشابه روند فوق، دوباره محاسبات را برای محاسبه $V(A)$ انجام می‌دهیم:

$$V(B) = \mathbb{E}[r] = 2p \times r_3 + (1-2p) \times r_4 = \frac{6p}{1+ap^2} + \frac{4(1-2p)}{1+ap^2} = \frac{4-2p}{1+ap^2}$$

$$V(C) = 3p \times r_2 + (1-3p) \times r_1 = \frac{6p}{1+ap^2} + \frac{(1-3p)}{1+ap^2} = \frac{1+3p}{1+ap^2}$$

$$V(A) = p \times V(B) + (1-p) \times V(C) = \frac{4p-2p^2}{1+ap^2} + \frac{(1-p)(1+3p)}{1+ap^2} = \frac{-5p^2+6p+1}{1+ap^2}$$

و مجدداً بر حسب p مشتق گرفته و مساوی صفر قرار می‌دهیم:

$$\frac{dV(A)}{dp} = 0 \Rightarrow \frac{d}{dp}\frac{-5p^2+6p+1}{1+ap^2} = \frac{(6-10p)(1+ap^2)-(2ap)(-5p^2+6p+1)}{(1+ap^2)^2}$$
$$= \frac{6-10p+6ap^2-10ap^3+10ap^3-12ap^2-2ap}{(1+p^2)^2} = \frac{6+(-10-2a)p-6ap^2}{(1+p^2)^2} = 0$$

همچنین می‌دانیم $p = \frac{1}{4}$ جواب معادله فوق است بنابراین با جایگذاری داریم:

$$p = \frac{1}{4} \Rightarrow 6 + (-10-2a) \times \frac{1}{4} - \frac{6a}{16} = 6 + \frac{-40-14a}{16} = 0 \Rightarrow 40+14a = 96 \Rightarrow a = \frac{56}{14} = 4$$

۳.

برای خارج شدن انتخاب‌های بازی از حالت احتمالاتی، لازم است تا توزیع‌های احتمالاتی به صورت one-hot شوند. به عبارت دیگر، مثلاً حالت A را در نظر بگیرید. برای آن که توزیع احتمال $(p, 1-p)$ به صورت قطعی شود، یا باید $p = 1$ باشد یا $1 - p = 1$. حالا توجه کنید که همین اتفاق برای حالت‌های B و C نیز باید برقرار باشد. حال توجه کنید که اگر $p = 1$ برقرار باشد، توزیع‌های B و C به صورت قطعی نخواهند بود و هنوز حالت احتمالاتی خواهند داشت (البته اصلاً توزیع احتمال معتبری هم نمی‌شوند) لذا باید به سراغ گزینه دیگر یعنی $1 - p = 1$ یا معادلاً $p = 0$ باید رفت. در این صورت دیده می‌شود که همه توزیع احتمالات به صورت قطعی در خواهند آمد.

با توجه به این که $p = 0$ شده است، در A و C همیشه اکشن سمت راست انتخاب می‌شود لذا لازم است تا بیشترین پاداش در راست‌ترین جایگاه قرار بگیرد. با در نظر گرفتن محدود بودن مجموع پاداش‌ها، بهترین حالت پخش پاداش‌ها به صورت $r_1 = 1$ و $r_2 = r_3 = r_4 = 0$ می‌باشد.