

- Vanilla-Momentum Gradient-Descent Observations:

Weight	Bias	Learning-Rate	Epochs	Vanilla-GD-Error	Momentum-GD-Error	Vanilla-Time	Momentum-time
-2	-2	10	700	$1e^{-31}$	$3.74e^{-32}$	0.0	0.015625
-4	-4	10	700	$2.5e^{-32}$	$6.16e^{-33}$	0.0	0.015625
-4	-4	15	1000	$7.7e^{-33}$	$7.7e^{-33}$	0.0	0.015625
-4	-4	20	1000	$6.5e^{-33}$	$3.58e^{-31}$	0.0	0.015625

- Conclusion:
- Here as you can see the table, here momentum gradient is taking more time than vanilla gradient but the error rate of momentum-gradient is less than vanilla gradient.
- Here the learning rate affects more to the error rate, than other parameters.
- As in the 3rd row, the error rate of momentum-gradient and vanilla-gradient becomes same here.
- And in the 4th row, the error rate of the vanilla is less than the momentum-gradient, so here the vanilla-gradient converges more than the momentum-gradient-descent.