Mean and Std of Attention and Feedforward Weights Across Layers of Neelectric/OLMo-2-1124-7B_MATH							
31	1.82e-06 (+/- 0.0183)	3.79e-05 (+/- 0.0194)	8.29e-06 (+/- 0.0192)	-2.77e-06 (+/- 0.0186)	-4.17e-06 (+/- 0.0161)	-9.28e-07 (+/- 0.0195)	9.13e-07 (+/- 0.0155)
30	3.35e-07 (+/- 0.0188)	1.7e-05 (+/- 0.0192)	5.66e-07 (+/- 0.0193)	4.17e-06 (+/- 0.019)	-4.71e-06 (+/- 0.0168)	-2.5e-06 (+/- 0.0194)	-9e-06 (+/- 0.0162)
29	-3.58e-06 (+/- 0.0189)	1.36e-05 (+/- 0.0188)	3.35e-06 (+/- 0.0193)	5.36e-08 (+/- 0.0192)	-3.04e-06 (+/- 0.0161)	4.83e-06 (+/- 0.0193)	4.02e-06 (+/- 0.0156)
- 28	-4.56e-07 (+/- 0.019)	7.93e-06 (+/- 0.0188)	-2.38e-06 (+/- 0.0193)	-1.66e-06 (+/- 0.0192)	7.23e-07 (+/- 0.0167)	6.02e-06 (+/- 0.0192)	-1.76e-06 (+/- 0.016)
27	3.19e-06 (+/- 0.019)	1.66e-05 (+/- 0.019)	8.61e-07 (+/- 0.0192)	-1.14e-06 (+/- 0.0192)	-9.18e-06 (+/- 0.0168)	-4.17e-06 (+/- 0.0192)	-6.97e-07 (+/- 0.0162)
26	-3.17e-06 (+/- 0.019)	1.66e-05 (+/- 0.0194)	-1.17e-06 (+/- 0.0194)	4.38e-06 (+/- 0.0193)	-7.53e-07 (+/- 0.017)	-3.04e-06 (+/- 0.0192)	-9.48e-06 (+/- 0.0162)
25	-8.72e-07 (+/- 0.019)	2.48e-05 (+/- 0.0198)	1.57e-06 (+/- 0.0193)	-1.07e-05 (+/- 0.0193)	-5.96e-06 (+/- 0.0167)	9.42e-07 (+/- 0.0193)	-2.31e-06 (+/- 0.0161)
24 -	-1.77e-06 (+/- 0.019)	1.47e-05 (+/- 0.02)	-9e-06 (+/- 0.0194)	-9.95e-06 (+/- 0.0193)	1.38e-06 (+/- 0.0172)	4.02e-06 (+/- 0.0193)	1.4e-06 (+/- 0.0166)
- 23	1.58e-06 (+/- 0.0192)	3.03e-05 (+/- 0.0203)	6.85e-06 (+/- 0.0193)	6.05e-06 (+/- 0.0192)	3.93e-06 (+/- 0.0167)	3.9e-06 (+/- 0.0193)	-2.15e-06 (+/- 0.0161)
22 -	-1.47e-06 (+/- 0.0192)	3.04e-05 (+/- 0.0205)	2.38e-06 (+/- 0.0194)	5.51e-06 (+/- 0.0193)	-3.04e-06 (+/- 0.0171)	4.71e-06 (+/- 0.0193)	-4.29e-06 (+/- 0.0167)
21	8.94e-07 (+/- 0.0192)	1.17e-05 (+/- 0.0205)	6.68e-06 (+/- 0.0194)	-1.31e-06 (+/- 0.0192)	3.04e-07 (+/- 0.0168)	-2.31e-06 (+/- 0.0194)	6.47e-06 (+/- 0.0161)
20	2.29e-06 (+/- 0.0192)	1.34e-05 (+/- 0.0206)	-1.24e-06 (+/- 0.0194)	-1.43e-06 (+/- 0.0193)	8.34e-06 (+/- 0.0167)	-1.86e-06 (+/- 0.0194)	1.8e-06 (+/- 0.0164)
19	9.5e-07 (+/- 0.0192)	1.72e-05 (+/- 0.0205)	-2.79e-06 (+/- 0.0194)	3.78e-06 (+/- 0.0192)	4.29e-06 (+/- 0.0168)	9.39e-07 (+/- 0.0194)	2.59e-07 (+/- 0.0165)
- 18	4.08e-07 (+/- 0.0192)	1.6e-05 (+/- 0.0208)	-2.31e-07 (+/- 0.0195)	5.48e-06 (+/- 0.0193)	2.79e-07 (+/- 0.0161)	7.71e-07 (+/- 0.0194)	8.23e-07 (+/- 0.0157)
17	8.34e-06 (+/- 0.0192)	1.32e-05 (+/- 0.0208)	1.25e-06 (+/- 0.0194)	1.37e-07 (+/- 0.0192)	3.9e-06 (+/- 0.0171)	-3.96e-06 (+/- 0.0194)	7.36e-06 (+/- 0.0168)
yer 16 -	-4.29e-06 (+/- 0.0192)	1.44e-05 (+/- 0.0206)	-2.2e-07 (+/- 0.0193)	1.94e-06 (+/- 0.019)	-2.1e-06 (+/- 0.0166)	-2.35e-06 (+/- 0.0194)	-5.04e-06 (+/- 0.0165)
La) 15 -	-2.92e-06 (+/- 0.019)	5.45e-06 (+/- 0.0203)	-4.17e-08 (+/- 0.0193)	-5.28e-06 (+/- 0.0193)	3.56e-06 (+/- 0.0168)	1.98e-06 (+/- 0.0193)	-4.51e-07 (+/- 0.0167)
14 -	-3.43e-07 (+/- 0.0189)	3.99e-06 (+/- 0.0203)	3.96e-06 (+/- 0.0193)	2.64e-07 (+/- 0.019)	3.8e-06 (+/- 0.0171)	-1.19e-06 (+/- 0.0192)	8.34e-07 (+/- 0.0171)
13	4.25e-07 (+/- 0.0189)	1.15e-06 (+/- 0.0205)	-5.16e-06 (+/- 0.0192)	1.55e-07 (+/- 0.019)	-1.71e-06 (+/- 0.0172)	3.7e-06 (+/- 0.0192)	-3.77e-06 (+/- 0.0171)
12	2.01e-06 (+/- 0.0189)	5.96e-06 (+/- 0.0205)	-2.34e-06 (+/- 0.019)	-6.5e-06 (+/- 0.0188)	-1.71e-06 (+/- 0.0166)	1.73e-06 (+/- 0.0192)	-7.82e-07 (+/- 0.0166)
11 -	4.27e-07 (+/- 0.0189)	-9.42e-07 (+/- 0.0208)	1.27e-05 (+/- 0.0194)	-3.38e-06 (+/- 0.0189)	-2.97e-06 (+/- 0.0167)	3.9e-06 (+/- 0.0192)	2.58e-06 (+/- 0.0167)
10	-6.03e-07 (+/- 0.0189)	1.98e-06 (+/- 0.0208)	4.86e-06 (+/- 0.0192)	6.94e-06 (+/- 0.0188)	3.84e-06 (+/- 0.0168)	-1.44e-06 (+/- 0.0192)	4.74e-06 (+/- 0.0167)
ი -	-2.74e-06 (+/- 0.0189)	1.22e-05 (+/- 0.0209)	-6.88e-06 (+/- 0.0189)	1.17e-05 (+/- 0.0188)	1.59e-06 (+/- 0.0145)	-4.2e-06 (+/- 0.0192)	-5.72e-06 (+/- 0.0143)
∞ -	-1.48e-06 (+/- 0.0189)	2.49e-06 (+/- 0.0204)	-6.38e-06 (+/- 0.0188)	1.36e-06 (+/- 0.0183)	1.97e-06 (+/- 0.0155)	-1.3e-06 (+/- 0.019)	-1.54e-06 (+/- 0.0154)
~ -	-5.69e-06 (+/- 0.019)	1.67e-05 (+/- 0.0208)	4.83e-06 (+/- 0.0189)	2.97e-06 (+/- 0.0186)	1.17e-06 (+/- 0.0157)	-1.11e-06 (+/- 0.0193)	2.12e-06 (+/- 0.0157)
9 -	3.65e-06 (+/- 0.019)	1.54e-05 (+/- 0.0214)	-3.05e-06 (+/- 0.019)	-3.93e-06 (+/- 0.0187)	-2.59e-06 (+/- 0.0154)	1.91e-06 (+/- 0.0193)	3.54e-07 (+/- 0.0153)
rv -	3.93e-06 (+/- 0.0192)	1.62e-05 (+/- 0.0208)	1.36e-05 (+/- 0.0189)	-1.23e-06 (+/- 0.0188)	-1.35e-06 (+/- 0.0154)	1.11e-06 (+/- 0.0193)	-4.98e-06 (+/- 0.0153)
4 -	7.45e-06 (+/- 0.0192)	3.08e-05 (+/- 0.0203)	-1.88e-06 (+/- 0.0186)	1.03e-06 (+/- 0.0183)	-5.19e-06 (+/- 0.0153)	6.47e-06 (+/- 0.0193)	4.02e-06 (+/- 0.0151)
ო -	8.79e-07 (+/- 0.0193)	3.53e-05 (+/- 0.0206)	-2.22e-06 (+/- 0.0188)	4.19e-07 (+/- 0.0188)	5.08e-08 (+/- 0.016)	3.58e-06 (+/- 0.0194)	-4.8e-06 (+/- 0.0159)
- 2	4.62e-06 (+/- 0.0193)	2.42e-05 (+/- 0.0199)	5.66e-06 (+/- 0.0188)	-8.12e-07 (+/- 0.0187)	-2.85e-06 (+/- 0.0156)	2.44e-06 (+/- 0.0192)	7.75e-06 (+/- 0.0153)
- ⊢	3.52e-06 (+/- 0.0195)	3.08e-05 (+/- 0.0204)	5.01e-06 (+/- 0.019)	1.67e-06 (+/- 0.0188)	3.62e-06 (+/- 0.0151)	2.62e-06 (+/- 0.0194)	-3.43e-06 (+/- 0.0153)
0 -	-1.18e-06 (+/- 0.019)	1.69e-05 (+/- 0.0199)	-7.3e-07 (+/- 0.0167)	-1.86e-06 (+/- 0.0164)	-2.97e-06 (+/- 0.0103)	9.91e-07 (+/- 0.0189)	-6.63e-07 (+/- 0.0104)
	down_proj Mean	gate_proj Mean	k_proj Mean	o_proj Mean	q_proj Mean	up_proj Mean	v_proj Mean

Submodule

1e-5