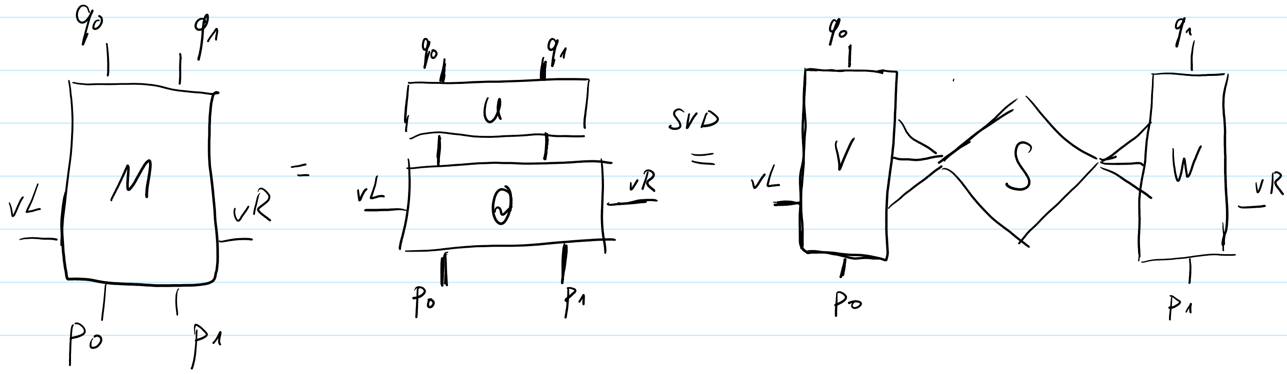
Calculating Euclidian Gradient

01 November 2021 21:09

We want to calculate Euclidian Gradient of Renyi entropy defined on the following tensor:



Cost function (Renyi Entropy):

$$\mathcal{J}^{\alpha}(U) = \frac{1}{1-\alpha} \log\left[\mathrm{tr}\left(S^{\alpha}\right)\right],$$

where Θ is assumed to be constant.

$$\frac{\partial \mathcal{J}(U)}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} = \frac{1}{1-\alpha} \frac{\partial}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} \left(\log\left[\mathrm{tr}(S)^{\alpha}\right] \right)$$

$$= \frac{1}{1-\alpha} \frac{1}{\mathrm{tr}(S^{\alpha})} \frac{\partial\, \mathrm{tr}(S^{\alpha})}{\partial U^{a_0 a_1}_{a_0^* a_1^*}}$$

To calculate gradients $\partial \mathcal{J}(U) / \partial U$, let's write tensors explicitly (using Einstein notation throughout, up / down indices do not indicate standard / dual spaces, they are just used to shorten the notation):

$$M^{q_0 q_1}_{vL\, vR\, p_0 p_1} = U^{q_0 q_1}_{q_0^* q_1^*}\, \Theta^{q_0^* q_1^*}_{vL\, vR\, p_0 p_1} = V^{q_0 q_0'}_{vL\, vL'\, p_0 p_0'}\, S^{vL'\, vR'\, p_0' p_1'}_{q_0' q_1'}\, W^{q_1 q_1'}_{vR\, vR'\, p_1 p_1'},$$

because V / W are left / right orthonormal we can write:

$$\left(V^{q_0 q_0''}_{vL\, vL''\, p_0 p_0''}\right)^{*} M^{q_0 q_1}_{vL\, vR\, p_0 p_1} \left(W^{q_1 q_1''}_{vR\, vR''\, p_1 p_1''}\right) =$$

$$= \underbrace{\left(V^{q_0 q_0''}_{v_L v_L'' p_0 p_0''}\right)^* V^{q_0 q_0'}_{v_L v_L' p_0 p_0'}}_{\delta_{q_0' q_0''}\,\delta_{v_L' v_L''}\,\delta_{p_0' p_0''}} \; S^{v_L' v_R' p_0' p_1'}_{q_0' q_1'} \; \underbrace{W^{q_1 q_1'}_{v_R v_R' p_1 p_1'} \left(W^{q_1 q_1''}_{v_R v_R'' p_1 p_1''}\right)}_{\delta_{q_1' q_1''}\,\delta_{v_R' v_R''}\,\delta_{p_1' p_1''}}$$

$$= S^{v_L'' v_R'' p_0'' p_1''}_{q_0'' q_1''}$$

From that we get:

$$\frac{\partial \operatorname{tr}(S^\alpha)}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} = \left[\text{because } S \text{ is diagonal}\right] = \frac{\partial}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} \left(S^{vv\,pp}_{qq}\right)^\alpha$$

here we don't sum over $\underline{v}, \underline{p}, \underline{q}$ but for each indices v, p, q we set $\underline{v}, \underline{p}, \underline{q}$ to same indices

$$= \alpha \left(S^{\underline{v}\underline{v}\,\underline{p}\underline{p}}_{\underline{q}\underline{q}}\right)^{\alpha-1} \frac{\partial}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} S^{vv\,pp}_{qq} =$$

$$= \alpha \left(S^{\underline{v}\underline{v}\,\underline{p}\underline{p}}_{\underline{q}\underline{q}}\right)^{\alpha-1} \frac{\partial}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} \left[\left(V^{q_0 q}_{v_L v\, p_0 p}\right)^* M^{q_0 q_1}_{v_L v_R\, p_0 p_1} W^{q_1 q}_{v_R v\, p_1 p}\right]$$

$$= \alpha \left(S^{\underline{v}\underline{v}\,\underline{p}\underline{p}}_{\underline{q}\underline{q}}\right)^{\alpha-1} \frac{\partial}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} \left[\left(V^{q_0 q}_{v_L v\, p_0 p}\right)^* U^{q_0 q_1}_{q_0^* q_1^*}\, \Theta^{q_0^* q_1^*}_{v_L v_R\, p_0 p_1} W^{q_1 q}_{v_R v\, p_1 p}\right]$$

$$= \alpha \left(S^{\underline{v}\underline{v}\,\underline{p}\underline{p}}_{\underline{q}\underline{q}}\right)^{\alpha-1} \delta_{a_0 q_0}\,\delta_{a_1 q_1}\,\delta_{a_0^* q_0^*}\,\delta_{a_1^* q_1^*} \left(V^{q_0 q}_{v_L v\, p_0 p}\right)^* \Theta^{q_0^* q_1^*}_{v_L v_R\, p_0 p_1} W^{q_1 q}_{v_R v\, p_1 p}$$

$$= \alpha \left(S^{\underline{v}\underline{v}\,\underline{p}\underline{p}}_{\underline{q}\underline{q}}\right)^{\alpha-1} \left(V^{a_0 q}_{v_L v\, p_0 p}\right)^* \Theta^{a_0^* a_1^*}_{v_L v_R\, p_0 p_1} W^{a_1 q}_{v_R v\, p_1 p}$$

So the partial derivatives of the cost function are:

$$\frac{\partial \zeta(U)}{\partial U^{a_0 a_1}_{a_0^* a_1^*}} = \frac{\alpha}{1-\alpha} \frac{1}{\operatorname{tr}(S^\alpha)} \left(S^{\underline{v}\underline{v}\,\underline{p}\underline{p}}_{\underline{q}\underline{q}}\right)^{\alpha-1} \left(V^{a_0 q}_{v_L v\, p_0 p}\right)^* \Theta^{a_0^* a_1^*}_{v_L v_R\, p_0 p_1} W^{a_1 q}_{v_R v\, p_1 p}$$