A. Proof of Theoretical Results

A.1. Proof of Theorem 2.1

Theorem. *Consider the optimization process in LoRA, where $\Delta W = AB$. Let A_0 denote the initial value of A, and suppose $A^* \in \mathbb{R}^{m\times r}$ and $B^* \in \mathbb{R}^{r\times n}$ are the optimal solutions for A and B in the LoRA optimization. Assume that both A_0 and A^* are full-rank matrices. Then, the low-rank update of W in LoRA can be expressed as a single-layer linear regression, formulated as:*

$$\Delta W = A^* B^* \approx A_0 B'^* \tag{11}$$

where B'^* has the same dimensions as B, i.e., $B'^* \in \mathbb{R}^{r\times n}$.

Proof. In summary, the proof of Theorem 2.1 is divided into two parts. In the first part we derive the optimal solution, in the second part we give the Expected Value of the approximation under the distribution of $N(0, 1)$.

Part I

Given the optimal solutions $A^* \in \mathbb{R}^{m\times r}$ and $B^* \in \mathbb{R}^{r\times n}$, the optimal update to W is given by:

$$\Delta W = A^* B^* \tag{12}$$

If there exists a matrix $C \in \mathbb{R}^{r\times r}$ such that $A^* = A_0 C$, substituting this into Equation 11 yields:

$$\Delta W = A^* B^* \approx A_0 (C B^*) \tag{13}$$

Since the composition of linear transformations is equivalent to a single linear transformation, CB^* can be reparameterized as B'^*. This demonstrates that the update represented in Equation 11 can be viewed as originating from a single-layer linear update.

Next, we derive the conditions under which such a C exists:

1. A^* and A_0 span the same subspace $\mathcal{S}$. If the two full-rank matrices A^* and A_0 span the same subspace $\mathcal{S}$, there exists an invertible matrix $C \in \mathbb{R}^{r\times r}$ such that $A^* = A_0 C$.

2. A^* and A_0 span different subspaces $\mathcal{S}' \neq \mathcal{S}$. If A^* and A_0 span different subspaces $\mathcal{S}'$ and $\mathcal{S}$, we can approximate C by solving the following optimization problem:

$$\min_C \|A^* - A_0 C\|_F^2 \tag{14}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Let the objective function be:

$$f(C) = \|A^* - A_0 C\|_F^2 \tag{15}$$

Taking the gradient of f with respect to C and setting it to zero, we obtain:

$$\frac{\partial f}{\partial C} = -2A_0^T (A^* - A_0 C) = 0 \tag{16}$$

$$A_0^T A_0 C = A_0^T A^* \tag{17}$$

Since $A_0^T A_0$ is invertible (as A_0 is full rank with rank r), we can solve for C:

$$C = (A_0^T A_0)^{-1} A_0^T A^* \tag{18}$$

This shows that C has an optimal closed-form solution. Furthermore, since A^* is the optimal solution for A in the LoRA optimization, the corresponding C is also optimal for updating W.

Part II

Next, we calculate the Expected Value of $\|A^* - A_0C^*\|_F^2$ when A^* and A_0 are under $N(0,1)$.

First, recall that the residual can be expressed as:

$$\begin{aligned} A^* - A_0C^* &= A^* - A_0(A_0^TA_0)^{-1}A_0^TA^* \\ &= (I-P)A^* \end{aligned}$$

where $P = A_0(A_0^TA_0)^{-1}A_0^T$ is the projection matrix onto the column space of A_0, $I-P$ is the projection onto the orthogonal complement of the column space of A_0. Therefore, the residual $A^* - A_0C^*$ is the projection of A^* onto the orthogonal complement of col(A_0).

Since A^* has entries from $N(0,1)$, its columns a_i^* are independent standard Gaussian vectors in $\mathbb{R}^m$. For each column a_i^*, the residual is:

$$r_i = (I-P)a_i^*$$

The squared Frobenius norm of the residual is:

$$\|A^* - A_0C^*\|_F^2 = \sum_{i=1}^{r} \|r_i\|_2^2$$

Since the columns are identically distributed, it suffices to compute:

$$E\left[\|r_i\|_2^2\right], \quad \text{for any } i = 1, \ldots, r$$

Because A^* and A_0 are independent and A_0 is fixed in each term when considering a_i^*, we condition on A_0. Let col(A_0) be the r-dimensional subspace of $\mathbb{R}^m$ spanned by the columns of A_0, col$(A_0)^\perp$ be its orthogonal complement, which has dimension $m-r$. The projection $(I-P)$ projects any vector onto col$(A_0)^\perp$.

Since $a_i^* \sim N(0, I_m)$, its variance in any direction is 1. The expected squared norm of its projection onto col$(A_0)^\perp$ is:

$$E\left[\|r_i\|_2^2 \mid A_0\right] = \sum_{j=1}^{m-r} E\left[(v_j^Ta_i^*)^2\right]$$

where $\{v_j\}$ is an orthonormal basis for col$(A_0)^\perp$.

Since $v_j^Ta_i^* \sim N(0,1)$, we have:

$$E\left[(v_j^Ta_i^*)^2\right] = 1$$

Thus:

$$E\left[\|r_i\|_2^2 \mid A_0\right] = m - r$$

Because this holds for each column a_i^*, the total expected squared norm is:

$$E\left[\|A^* - A_0C^*\|_F^2 \mid A_0\right] = r \times (m-r)$$

Since $E\left[\|A^* - A_0C^*\|_F^2 \mid A_0\right]$ does not depend on A_0 (the result is the same for any full-rank A_0), we have:

$$E\left[\|A^* - A_0C^*\|_F^2\right] = r \times (m-r)$$

Therefore, the expected value is as above.

$\square$