

COMP/LING 445/645

Problem Set 4

This problem set consists only of questions involving mathematics or English or or a combination of the two. Please put your answers in an **Answer** section like in the example below. You can find more information about L^AT_EX here <https://www.latex-project.org/>.

Once you have answered the question, please compile your copy of this L^AT_EX document into a PDF and submit (i) the compiled PDF (ii) the raw L^AT_EX file and (iii) your **lastname-firstname.clj** file via email to *both* timothy.odonnell@mcgill.ca and savanna.willerton@mail.mcgill.ca. The problem is set is due before 16:05 on Monday, November 25, 2019.

Problem 0: This is an example question using some fake math like this $L = \sum_0^\infty \mathcal{G}\delta_x$.

Answer 0: Put your answer right under the question like this $L = \sum_0^\infty \mathcal{G}\delta_x$.

Problem 1: In class we gave the following equation for the bigram probability of a sequence of words $W^{(1)}, \dots, W^{(k)}$:

$$\Pr(W^{(1)}, \dots, W^{(k)}) = \prod_i^k \Pr(W^{(i)} | W^{(i-1)} = w^{(i-1)}) \quad (1)$$

Using this formula, give an expression for the bigram probability of the sentence *abab*, where each character is treated as a word. Try to simplify the formula as much as possible.

Answer 1: Expanding the above equation for $k = 4$, we get:

$$\begin{aligned} \Pr(\times abab \times) &= \Pr(a | W^{(0)} = \times) \times \Pr(b | W^{(1)} = a) \times \Pr(a | W^{(2)} = b) \times \Pr(b | W^{(3)} = a) \\ &= \Pr(a | W^{(0)} = \times) \times \Pr(b | W^{(1)} = a)^2 \times \Pr(a | W^{(2)} = b) \end{aligned}$$

Problem 2: Suppose that there are two possible symbols/words in our language, a and b . There are three conditional distributions in the bigram model for this language, $\Pr(W^{(i)} | W^{(i-1)} = a)$, $\Pr(W^{(i)} | W^{(i-1)} = b)$, and $\Pr(W^{(i)} | W^{(i-1)} = \times)$. These conditional distributions are associated with the parameter vectors $\vec{\theta}_a$, $\vec{\theta}_b$, and $\vec{\theta}_\times$, respectively (these parameter vectors were implicit in the previous problem). For the current problem, we will assume that these parameters are fixed.

Suppose that we are given a sentence $W^{(1)}, \dots, W^{(k)}$. We will use the notation $n_{x \rightarrow y}$ to denote the number of times that the symbol y occurs immediately following the symbol x in the sentence. For example, $n_{a \rightarrow a}$ counts the number of times that symbol a occurs immediately following the symbol a . Using Equation 1, give an expression for the probability of a sentence in our language:

$$\Pr(W^{(1)}, \dots, W^{(k)} | \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times)$$

The expression should make use of the $n_{x \rightarrow y}$ notation defined above.

Hint: the expression should be analogous to the formula that we found for the likelihood of a corpus under a bag of words model.

Answer 2: Given a sentence $W^{(1)}, \dots, W^{(k)}$ indexed by i , we know from equation (1) that the probabilities that some word $W^{(i)}$ follows one of \times , a or b are given by the conditional probabilities stated above, which are associated with the parameter vectors $\vec{\theta}_a$, $\vec{\theta}_b$, and $\vec{\theta}_\times$. We also know that the probability that a sentence S is given by some distribution $\vec{\theta}$ is:

$$\Pr(S = w^{(1)}, \dots, w^{(k)} | \vec{\theta}) = \prod_{w \in S} \theta_w^{n_w} \quad (2)$$

Where n_w = count of w in S .

Combining the conditional distributions of the bigram BOW model and the probability given above, we can deduce an expression for the probability of a sentence by applying the chain rule:

$$\Pr(S = W^{(1)}, \dots, W^{(k)} | \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times) = \Pr(S_a | \vec{\theta}_a) \times \Pr(S_b | \vec{\theta}_b) \times \Pr(S_\times | \vec{\theta}_\times)$$

Where S_a, S_b , and S_\times represent subsets of the vocabulary of S where the preceding word is a , b and \times respectively. Expanding this knowing (2) yields:

$$\prod_{w \in S_\times} \theta_{\times \rightarrow w}^{n_{\times \rightarrow w}} \times \prod_{w \in S_a} \theta_{a \rightarrow w}^{n_{a \rightarrow w}} \times \prod_{w \in S_b} \theta_{b \rightarrow w}^{n_{b \rightarrow w}} \quad (3)$$

Problem 3: Assume the parameter vectors in our bigram model have the following values:

$$\vec{\theta}_a = (0.7, 0.3)$$

$$\vec{\theta}_b = (0.2, 0.8)$$

$$\vec{\theta}_\times = (0.5, 0.5)$$

The first vector indicates that if the current symbol a , there is probability 0.7 of transitioning to the symbol a , and probability 0.3 of transitioning to the symbol b . Using your answer to the previous problem and these parameter values, calculate the probability of the string $aabab$.

Answer 3: $\Pr(\times aabab \times) = (0.5)(0.7)(0.3)(0.2)(0.3) = (0.5)(0.7)(0.3)^2(0.2) = 0.0063$

Problem 4: In Problem 2, you found an expression for the bigram probability of a sentence in our language, which contains the symbols a and b . In that problem, we assumed that there were fixed parameter vectors $\vec{\theta}$ associated with each conditional distribution. In this problem, we will consider the setting in which we have uncertainty about the value of these parameters.

In class, we used the Dirichlet distribution to define a prior distribution over parameter vectors:

$$\vec{\theta}_c | \vec{\alpha} \sim \text{Dirichlet}(\vec{\alpha}) \quad (4)$$

$$w^{(i)} | w^{(i-1)} \sim \text{categorical}(\vec{\theta}_{w^{(i-1)}}) \quad (5)$$

$$w^{(1)} \sim \text{categorical}(\vec{\theta}_\times) \quad (6)$$

Suppose that we have a sentence $S = W^{(1)}, \dots, W^{(k)}$. Given an expression for the joint probability $\Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times | \vec{\alpha})$ using the definitions of Dirichlet distributions and likelihoods we defined in class.

Answer 4: The joint probability of a hierarchical model $\Pr(S, \vec{\theta} | \vec{\alpha})$ can be found by multiplying the likelihood of a sentence given some θ , $\Pr(S | \vec{\theta})$ by a Dirichlet prior over some θ with pseudo-counts α , $\Pr(\vec{\theta} | \vec{\alpha})$. We know the likelihood of some sentence in our model by equation (3). We can find the Dirichlet distributions representing the bigram model parameters by using:

$$\Pr(\vec{\theta} | \vec{\alpha}) = \frac{1}{B(\vec{\alpha})} \prod_{i=1}^K \theta_i^{\alpha_i-1} = \frac{\Gamma(\sum_{i=1}^K \alpha_i)}{\prod_{i=1}^K \Gamma(\alpha_i)} \prod_{i=1}^K \theta_i^{\alpha_i-1}$$

Where $\vec{\theta}$ and $\vec{\alpha}$ are indexed by i . The bigram model parameters are:

$$\Pr(\vec{\theta}_a | \vec{\alpha}) = \frac{1}{B(\vec{\alpha})} \prod_{w \in S_a} \theta_{a \rightarrow w}^{\alpha_w-1} = \frac{\Gamma(\sum_{w \in S_a} \alpha_w)}{\prod_{w \in S_a} \Gamma(\alpha_w)} \prod_{w \in S_a} \theta_{a \rightarrow w}^{\alpha_w-1} \quad (7)$$

$$\Pr(\vec{\theta}_b | \vec{\alpha}) = \frac{1}{B(\vec{\alpha})} \prod_{w \in S_b} \theta_{b \rightarrow w}^{\alpha_w-1} = \frac{\Gamma(\sum_{w \in S_b} \alpha_w)}{\prod_{w \in S_b} \Gamma(\alpha_w)} \prod_{w \in S_b} \theta_{b \rightarrow w}^{\alpha_w-1} \quad (8)$$

$$\Pr(\vec{\theta}_x | \vec{\alpha}) = \frac{1}{B(\vec{\alpha})} \prod_{w \in S_x} \theta_{x \rightarrow w}^{\alpha_w-1} = \frac{\Gamma(\sum_{w \in S_x} \alpha_w)}{\prod_{w \in S_x} \Gamma(\alpha_w)} \prod_{w \in S_x} \theta_{x \rightarrow w}^{\alpha_w-1} \quad (9)$$

Now by multiplying and using chain rule, we get the formula for the joint probability of our model:

$$\begin{aligned} \Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_x | \vec{\alpha}) &= \Pr(S, \vec{\theta}_a | \vec{\alpha}) \Pr(S, \vec{\theta}_b | \vec{\alpha}) \Pr(S, \vec{\theta}_x | \vec{\alpha}) \\ &= \Pr(S | \vec{\theta}_a) \Pr(\vec{\theta}_a | \vec{\alpha}) \times \Pr(S | \vec{\theta}_b) \Pr(\vec{\theta}_b | \vec{\alpha}) \times \Pr(S | \vec{\theta}_x) \Pr(\vec{\theta}_x | \vec{\alpha}) \\ &= \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \times \frac{\Gamma(\sum_{w \in S_b} \alpha_w)}{\prod_{w \in S_b} \Gamma(\alpha_w)} \prod_{w \in S_b} \theta_{b \rightarrow w}^{[n_{b \rightarrow w} + \alpha_w] - 1} \times \frac{\Gamma(\sum_{w \in S_x} \alpha_w)}{\prod_{w \in S_x} \Gamma(\alpha_w)} \prod_{w \in S_x} \theta_{x \rightarrow w}^{[n_{x \rightarrow w} + \alpha_w] - 1} \\ &= \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \times \left(\prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_b} \theta_{b \rightarrow w}^{[n_{b \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_x} \theta_{x \rightarrow w}^{[n_{x \rightarrow w} + \alpha_w] - 1} \right) \end{aligned} \quad (10)$$

Problem 5: In the previous problem, you found a formula for the joint probability of a sentence and a set of bigram model parameters. Using this, give a formula for the marginal probability of the sentence $\Pr(S | \vec{\alpha})$.

Hint: The formula should be analogous to the formula derived in class for marginal probability of a corpus under a bag of words model. Whereas before there was only a single parameter vector $\vec{\theta}$, now there are three parameters that need to be marginalized away. Otherwise the calculation will be similar.

Answer 5: First we can note that the following expression holds:

$$\int_{\Theta} \prod_{i=1}^K \theta_i^{x_i-1} = \frac{\prod_{i=1}^K \Gamma(x_i)}{\Gamma(\sum_{i=1}^K x_i)} \quad (11)$$

$$\begin{aligned}
\Pr(\mathbf{S}|\vec{\alpha}) &= \int_{\Theta_a} \int_{\Theta_b} \int_{\Theta_{\times}} \Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_{\times} | \vec{\alpha}) \\
&= \int_{\Theta_a} \int_{\Theta_b} \int_{\Theta_{\times}} \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \times \left(\prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_b} \theta_{b \rightarrow w}^{[n_{b \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_{\times}} \theta_{\times \rightarrow w}^{[n_{\times \rightarrow w} + \alpha_w] - 1} \right) \\
&= \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \frac{\prod_{w \in S_{\times}} \Gamma([n_{\times \rightarrow w} + \alpha_w])}{\Gamma(\sum_{w \in S_{\times}} [n_{\times \rightarrow w} + \alpha_w])} \int_{\Theta_a} \int_{\Theta_b} \times \left(\prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_b} \theta_{b \rightarrow w}^{[n_{b \rightarrow w} + \alpha_w] - 1} \right) \\
&= \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \frac{\prod_{w \in S_{\times}} \Gamma([n_{\times \rightarrow w} + \alpha_w])}{\Gamma(\sum_{w \in S_{\times}} [n_{\times \rightarrow w} + \alpha_w])} \frac{\prod_{w \in S_b} \Gamma([n_{b \rightarrow w} + \alpha_w])}{\Gamma(\sum_{w \in S_b} [n_{b \rightarrow w} + \alpha_w])} \int_{\Theta_a} \times \left(\prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \right) \\
&= \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \frac{\prod_{w \in S_{\times}} \Gamma([n_{\times \rightarrow w} + \alpha_w])}{\Gamma(\sum_{w \in S_{\times}} [n_{\times \rightarrow w} + \alpha_w])} \frac{\prod_{w \in S_b} \Gamma([n_{b \rightarrow w} + \alpha_w])}{\Gamma(\sum_{w \in S_b} [n_{b \rightarrow w} + \alpha_w])} \frac{\prod_{w \in S_a} \Gamma([n_{a \rightarrow w} + \alpha_w])}{\Gamma(\sum_{w \in S_a} [n_{a \rightarrow w} + \alpha_w])}
\end{aligned}$$

Problem 6: Let us assume that the parameters of the Dirichlet distribution are $\vec{\alpha} = (1, 1)$. Using your solution to the previous problem, compute the marginal probability of the string $aabab$. The formula that you compute should contain the gamma function $\Gamma(\cdot)$. Using the properties of the gamma function discussed in class (i.e., its relationship to the factorial) or an online calculator, compute a numerical value for this expression.

Answer 6: We define $\Gamma(n) = (n-1)!$ in the following equation.
From the string $\times aabab$, we get n values:

$$\begin{aligned}
n_{a \rightarrow a} &= 1 \\
n_{a \rightarrow b} &= 2 \\
n_{b \rightarrow a} &= 1 \\
n_{\times \rightarrow a} &= 1
\end{aligned}$$

And subsets:

$$\begin{aligned}
S_a &= \{a, b\} \\
S_b &= \{a\} \\
S_{\times} &= \{a\}
\end{aligned}$$

Plugging them into (12) we get:

$$\begin{aligned}
\Pr(\times aabab | (1, 1)) &= \frac{\Gamma(5)}{\Gamma(1)^5} \frac{\Gamma(2)}{\Gamma(2)} \frac{\Gamma(2)}{\Gamma(2)} \frac{\Gamma(2)\Gamma(3)}{\Gamma(5)} \\
&= (24)(1)(1)(2/24) \\
&= 2
\end{aligned}$$

Problem 7: Suppose that we have observed a sentence $S = W^{(1)}, \dots, W^{(k)}$. Find an expression for the posterior distribution over the model parameters, $\Pr(\vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_{\times} | S, \vec{\alpha})$.

Hint: Use the joint probability that you computed in Problem 4 and Bayes' rule. The solution should be analogous to the posterior probability for the bag of words model.

Answer 7: From Q4 we know that

$$\Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times | \vec{\alpha}) = \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \times \left(\prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_b} \theta_{b \rightarrow w}^{[n_{b \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_\times} \theta_{\times \rightarrow w}^{[n_{\times \rightarrow w} + \alpha_w] - 1} \right)$$

Bayes' rule is defined by:

$$\begin{aligned} \Pr(\vec{\theta} | S, \vec{\alpha}) &= \frac{\Pr(S, \vec{\theta} | \vec{\alpha})}{\Pr(S | \vec{\alpha})} \\ \Pr(\vec{\theta} | S, \vec{\alpha}) &= \frac{\Gamma(\sum_{w \in V} [n_w + \alpha_w])}{\prod_{w \in V} \Gamma([n_w + \alpha_w])} \prod_{w \in V} \theta_w^{[n_w + \alpha_w] - 1} \end{aligned}$$

Applying Bayes' rule to the equations in Q4 and Q5, we get:

$$\begin{aligned} \Pr(\vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times | S, \vec{\alpha}) &= \frac{\Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times | \vec{\alpha})}{\Pr(S | \vec{\alpha})} \\ &= \frac{\Gamma(\sum_{w \in V} \alpha_w)}{\prod_{w \in V} \Gamma(\alpha_w)} \times \left(\prod_{w \in S_a} \theta_{a \rightarrow w}^{[n_{a \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_b} \theta_{b \rightarrow w}^{[n_{b \rightarrow w} + \alpha_w] - 1} \prod_{w \in S_\times} \theta_{\times \rightarrow w}^{[n_{\times \rightarrow w} + \alpha_w] - 1} \right) \times \\ &\quad \frac{\prod_{w \in V} \Gamma(\alpha_w)}{\Gamma(\sum_{w \in V} \alpha_w)} \frac{\Gamma(\sum_{w \in S_\times} [n_{\times \rightarrow w} + \alpha_w])}{\prod_{w \in S_\times} \Gamma([n_{\times \rightarrow w} + \alpha_w])} \frac{\Gamma(\sum_{w \in S_b} [n_{b \rightarrow w} + \alpha_w])}{\prod_{w \in S_b} \Gamma([n_{b \rightarrow w} + \alpha_w])} \frac{\Gamma(\sum_{w \in S_a} [n_{a \rightarrow w} + \alpha_w])}{\prod_{w \in S_a} \Gamma([n_{a \rightarrow w} + \alpha_w])} \end{aligned}$$

Problem 8: Using your formula for the posterior distribution and setting $\vec{\alpha} = (1, 1)$, calculate the posterior distribution over parameters given the sentence *aabab*. There should be an easy way to interpret the posterior distribution, and how it was derived from the observed sentence. What is this interpretation?

Answer 8: From previous questions, we know :

$$\begin{aligned} \Pr(S | \vec{\alpha}) &= \Pr(\times aabab | (1, 1)) = 2 \\ \Pr(S, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times | \vec{\alpha}) &= \Pr(\times aabab, \vec{\theta}_a, \vec{\theta}_b, \vec{\theta}_\times | (1, 1)) \\ &= \frac{\Gamma(1)^5}{\Gamma(5)} \theta_{a \rightarrow a} \theta_{a \rightarrow b}^2 \times \theta_{b \rightarrow a} \times \theta_{\times \rightarrow a} \end{aligned}$$

Dividing we get:

$$\begin{aligned} &= \frac{1}{2} \frac{1}{24} \theta_{a \rightarrow a} \theta_{a \rightarrow b}^2 \times \theta_{b \rightarrow a} \times \theta_{\times \rightarrow a} \\ &= \left(\frac{1}{12}\right) \theta_{a \rightarrow a} \theta_{a \rightarrow b}^2 \times (1) \theta_{b \rightarrow a} \times (1) \theta_{\times \rightarrow a} \end{aligned}$$

The posterior distribution allows us to average over all of our posterior values to help predict the next step in case we don't want to commit to a specific model. The algorithm renormalizes after the addition of each term, and the averages are returned based on the history of the corpus.

Problem 9: Consider the language $L = \{a^*ba^*\}$, that is, the language consisting of some number of a's, followed by a single b, followed by some number of a's. Show that this language is not strictly 2-local.

Hint: use n -Local Suffix Substitution Closure (n -SSC).

Answer 9: Let $s1 = \times aabaaa$ where $u_1 = \times aab$, $x = a$ and $v_1 = aa$ and $s2 = \times aaabaa$ where $u_2 = \times aa$, $x = a$ and $v_2 = baa$

By n-SSC, we know that if a language is n-local and a pivot x is defined by a string of $n - 1$ length, then the sentence $u_1 x v_2 \in L$.

$u_1 x v_2 = \times aababaa \notin L$. L is not strictly 2-local.

Problem 10: Consider the language $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$, that is, the language consisting of n a's followed by m b's followed by n c's followed by m d's where n and m are natural numbers. Show that this language is not strictly 2-local.

Hint: use the same property as in the problem above.

Answer 10: Let $s1 = \times abcd$ where $u_1 = \times ab$, $x = c$ and $v_1 = d$ and $s2 = \times abcccd$ where $u_2 = \times ab$, $x = c$ and $v_2 = cd$

By n-SSC, we know that if a language is n-local and a pivot x is defined by a string of $n - 1$ length, then the sentence $u_1 x v_2 \in L$.

$u_1 x v_2 = \times abcccd$. We can see that for a , $n = 1$ but for c , $n = 2$. In L , a and c have the same exponent n but not in our new sentence, so $u_1 x v_2 \notin L$. L is not strictly 2-local.

Problem 11: Show that the language $L = \{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$ is not strictly k -local, for any value of k .

Answer 11: Let $s1 = \times a^n b^m c^n d^m$ where $u_1 = \times a^n b^m$, $x = c^{n-1}$ and $v_1 = cd^m$ and $s2 = \times a^{n-1} b^m c^{n-1} d^m$ where $u_2 = \times a^{n-1} b^m$, $x = c^{n-1}$ and $v_2 = d^m$

By n-SSC, we know that if a language is n-local and a pivot x is defined by a string of $n - 1$ length, then the sentence $u_1 x v_2 \in L$.

$u_1 x v_2 = \times a^n b^m c^{n-1} d^m$. We can clearly see that $n \neq n - 1$ so the new sentence does not satisfy the equation. $u_1 x v_2 \notin L$. This is true for any value of n . L is not strictly k -local for any value of k .

Problem 12: In class we proved that $k\text{-SSC}(L) \implies L \in \text{SL}_k$. In other words, if a language satisfies k -Local Suffix Substitution Closure, then it is k -strictly local.

Use this theorem to prove that k -strictly local languages are closed under intersection. More precisely, prove that if $L_1 \in \text{SL}_k$ and $L_2 \in \text{SL}_k$, then $L_1 \cap L_2 \in \text{SL}_k$.

Answer 12: Given the assumption that $L_1 \in \text{SL}_k$ and $L_2 \in \text{SL}_k$ both satisfy $n\text{-SSC}(L)$, we have by definition of $n\text{-SSC}(L)$ that for every $u_1, u_2, v_1, v_2 \in V_1^*$ and $x \in V_1^{n-1}$, if $u_1 x v_1 \in L_1$ and $u_2 x v_2 \in L_1$, then $u_1 x v_2 \in L_1$ also. The same follows for L_2 .

Taking the intersection of the values $u_1 x v_2$ of V_1^* and V_2^* , we get:

$$u_1 x v_1 \in L_1 \cap L_2 \wedge u_2 x v_2 \in L_1 \cap L_2 \implies u_1 x v_2 \in L_1 \cap L_2.$$

So, we see that if $L_1 \in \text{SL}_k$ and $L_2 \in \text{SL}_k$, then $L_1 \cap L_2 \in \text{SL}_k$.