18.413: Error-Correcting Codes Lab February 26, 2004

Lecture 7

Lecturer: Daniel A. Spielman

To begin, let me point out that there was a typo in the lecture notes from last lecture. Lemma 6.4.1 should have said:

Lemma 7.0.1.

$$P^{ext}\left[X_1 = a_1 | Y_2 Y_3 = b_2 b_3\right] = \sum_{a_2 : (a_1, a_2) \in \mathcal{C}_{12}} P\left[X_2 = a_2 | X_1 = a_1\right] P^{ext}\left[X_2 = a_2 | Y_2 = b_2\right] P^{ext}\left[X_2 = a_2 | Y_3 = b_3\right].$$

7.1 Markov Property

Last lecture, we considered three variables X_1, X_2 and X_3 chosen uniformly from those that satisfy $(X_1, X_2) \in \mathcal{C}_{12} \subseteq A_1 \times A_2$ and and $(X_2, X_3) \in \mathcal{C}_{23} \subseteq A_2 \times A_3$.

We claimed that the variables (X_1, X_2, X_3) then satisfy what the book calls the "Markov" property. That is, for all a_1, a_2, a_3,

$$\mathrm{P}\left[X_1 X_3 = a_1 a_3 | X_2 = a_2\right] = \mathrm{P}\left[X_1 = a_1 | X_2 = a_2\right] \mathrm{P}\left[X_3 = a_3 | X_2 = a_2\right].$$

I'll now sketch a proof. It basically follows by by recalling the definition of the probability of an even conditioned on $X_2 = a_2$. We first note that the set of choices for (X_1, X_2, X_3) given that $X_2 = a_2$ is

$$S_{a_2} \stackrel{\text{def}}{=} \left\{(X_1, a_2, X_3) : (X_1, a_2) \in \mathcal{C}_{12} \text{ and } (a_2, X_3) \in \mathcal{C}_{23}\right\}.$$

Conditioning on $X_2 = a_2$, we obtain a sample chosen uniformly from S_{a_2}. Thus, for $(a_1, a_2, a_3) \in S_{a_2}$,

$$\mathrm{P}\left[(X_1, X_2, X_3) = (a_1, a_2, a_3) | X_2 = a_2\right] = \frac{1}{|S_{a_2}|}.$$

Note that

$$|S_{a_2}| = |\{a_1 : (a_1, a_2) \in \mathcal{C}_{12}\}| \, |\{a_3 : (a_2, a_3) \in \mathcal{C}_{23}\}|$$

The claim now follows from observing that

$$\mathrm{P}\left[(X_1, X_2) = (a_1, a_2) | X_2 = a_2\right] = \frac{|\{a_3 : (a_2, a_3) \in \mathcal{C}_{23}\}|}{|S_{a_2}|} = \frac{1}{|\{a_1 : (a_1, a_2) \in \mathcal{C}_{12}\}|}$$

and

$$\mathrm{P}\left[(X_2, X_3) = (a_2, a_3) | X_2 = a_2\right] = \frac{|\{a_1 : (a_1, a_2) \in \mathcal{C}_{12}\}|}{|S_{a_2}|} = \frac{1}{|\{a_3 : (a_2, a_3) \in \mathcal{C}_{23}\}|}.$$

7.2 Simplifying Probability Computation

First, lets establish that the fundamental quantities we are interested in have the form

$$\mathrm{P}\left[X_i = a_i | E\right],$$

where E is some event, usually a union of the observed variables. We will typically want these values for all a_i, so we really want a vector

$$\left(\mathrm{P}\left[X_i = a_1|E\right], \mathrm{P}\left[X_i = a_2|E\right], \ldots, \mathrm{P}\left[X_i = a_n|E\right], \right),$$

where $a_1, \ldots, a_n$ are the symbols in the alphabet A_i. We will denote such a vector by

$$\vec{\mathrm{P}}\left[X_i|E\right].$$

Using this notation, and letting $\odot$ denote componentwise product ($(a,b) \odot (c,d) = (ac, bd)$), we have

$$\vec{\mathrm{P}}^{post}\left[X_i|E\right] = \vec{\mathrm{P}}^{prior}\left[X_i\right] \odot \vec{\mathrm{P}}^{ext}\left[X_i|E\right]$$

Before returning to our probability computations for (X_1, X_2, X_3), I'll also introduce the simpler notation $\mathrm{P}^{ext}\left[X_i = a_i|Y_i\right]$ for $\mathrm{P}^{ext}\left[X_i = a_i|Y_i = b_i\right]$. We will use this notation whenever b_i is fixed throughout our computation, which it generally is as it is what was received.

We then have, from Lemma 6.2.1,

$$\vec{\mathrm{P}}^{post}\left[X_1|Y_1Y_2Y_3\right] = \vec{\mathrm{P}}^{prior}\left[X_1\right] \odot \vec{\mathrm{P}}^{ext}\left[X_1|Y_1\right] \odot \vec{\mathrm{P}}^{ext}\left[X_1|Y_2Y_3\right],$$

and, from Lemma 7.0.1,

$$\mathrm{P}^{ext}\left[X_1 = a_1|Y_2Y_3\right] = \sum_{a_2:(a_1,a_2)\in\mathcal{C}_{12}} \mathrm{P}\left[X_2 = a_2|X_1 = a_1\right] \mathrm{P}^{ext}\left[X_2 = a_2|Y_2\right] \mathrm{P}^{ext}\left[X_2 = a_2|Y_3\right].$$

Using these formulas, we go through the following steps to compute $\vec{\mathrm{P}}^{post}\left[X_1|Y_1Y_2Y_3\right]$.

1. Compute $\vec{\mathrm{P}}^{ext}\left[X_2|Y_3\right]$. This computation only depends upon Y_3, and comes from the formula:
$$\mathrm{P}^{ext}\left[X_2 = a_2|Y_3\right] \sim \sum_{a_3:(a_2,a_3)\in\mathcal{C}_{23}} \mathrm{P}\left[X_3 = a_3|X_2 = a_2\right] \mathrm{P}\left[Y_3|X_3 = a_3\right].$$
2. Compute $\vec{\mathrm{P}}^{ext}\left[X_2 = |Y_2\right]$, and compute
$$\vec{\mathrm{P}}^{ext}\left[X_2 = |Y_2\right] \odot \mathrm{P}^{ext}\left[X_2 = a_2|Y_3\right].$$
3. For each a_1, compute
$$\vec{\mathrm{P}}\left[X_2|X_1 = a_1\right] \odot \vec{\mathrm{P}}^{ext}\left[X_2|Y_2\right] \odot \vec{\mathrm{P}}^{ext}\left[X_2|Y_3\right],$$
and then sum the resulting vector.
4. Take the output of the previous step, and $\odot$ product it with $\vec{\mathrm{P}}^{prior}\left[X_1\right] \odot \vec{\mathrm{P}}^{ext}\left[X_1|Y_1\right]$.

If you look at the flow of this computation, it can be understood as a vector being passed from X_3 to X_2 between steps 1 and 2, and a vector begin passed from X_2 to X_1 between steps 3 and 4.