نظریه اطلاعات شبکه، ترم بهار ۱۳۹۴-۱۳۹۳

مدرس: امین زاده گوهری

جلسه ۲

مرور بر مفاهیم اصلی نظریه اطلاعات نقطه به نقطه: زنجیره های مارکف

از آنجایی که زنجیره های مارکف در مسائل زیادی در تئوری اطلاعات، خصوصا هنگام نوشتن قسمت وارون، ظاهر میشوند بصورت مختصر بر آنها مروری خواهیم داشت.

تعریف ۱ میگوییم که سه متغیر تصادفی X, Y, Z زنجیره مارکف تشکیل میدهند و آن را با نماد $X - Y - Z$ و یا $Z \to Y \to X$ و یا $X \to Y \to Z$ نمایش میدهیم اگر و فقط اگر

$$p(x, y, z) = p(y)p(x|y)p(z|y)$$

برقرار باشد. این رابطه در مورد توزیع مشترک را میتوان بشکل

$$p(x, y, z) = p(x, y)p(z|y)$$

یا

$$p(x, y, z) = p(y, z)p(x|y)$$

نیز نوشت.

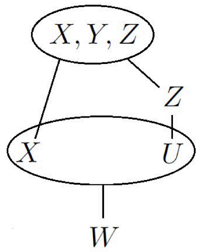
زنجیره مارکف به این معنا است که $I(X; Z|Y) = 0$. بصورت شهودی اگر روی بخش Y مشروط بکنیم، همانند این است که متغیر وسط زنجیر را از بین برده ایم؛ در نتیجه ارتباط میان X و Z قطع شده و از هم مستقل میشوند. در نتیجه به شرط دانستن Y این دو متغیر از هم مستقل هستند.

تمرین ۲ اگر یک زنجیره مارکف به شکل $XY - Z - W$ داشته باشیم، آنوقت زنجیره $X - Z - W$ هم برقرار است. یعنی میتوان از دو سمت زنجیره یک متغیر را حذف کرد. همچنین رابطه $X - YZ - W$ هم برقرار است. یعنی میتوان متغیرها را از دو طرف زنجیره به وسط آورد. اینها عملیاتهای بسیار پرکاربردی هستند.

زمانی که بیش از سه متغیر داشته باشیم انواع مختلفی از ارتباطات میان متغیرها قابل تصور است.

مثال ۳ برای چهار متغیر X, Y, Z, U رابطه $X - Y - Z - U$ به این معنا است که

$$\begin{aligned} p(x, y, z, u) &= p(x)p(y|x)p(z|y)p(u|z) \\ &= p(u)p(z|u)p(y|z)p(x|y) \end{aligned}$$



شکل ۱: درخت مارکف مربوط به توزیع $p(x, y, z, u, w) = p(x)p(y|x)p(z|x, y)p(u|z)p(w|u, x)$ در مثال ۴.

در اینجا میبینیم که اگر مثلا Y، Z و یا هر دو آنها را از زنجیره حذف کنیم، آنوقت ارتباط میان X و U قطع میشود، پس داریم:

$$I(X; U|Y) = I(X; U|Z) = I(X; U|Y, Z) = 0.$$

مشابها اگر X و Z را از زنجیر $X - Y - Z - U$ حذف کنیم، ارتباط میان U و Y قطع میشود. پس

$$I(Y; U|XZ) = 0.$$

بنابراین با انواع برش های مختلف میتوان به روابط مختلف رسید.

مثال ۴ فرض کنید که با متغیر X شروع کرده ایم. سپس متغیر Y از روی X؛ متغیر Z با عبور X, Y از یک کانال؛ متغیر U را با عبور Z از یک کانال و نهایتا متغیر W را از روی X و U ساخته ایم. یا به عبارت دیگر

$$p(x, y, z, u, w) = p(x)p(y|x)p(z|x, y)p(u|z)p(w|u, x).$$

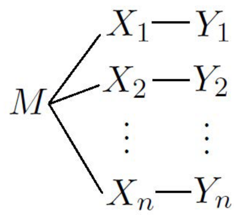
در اینجا زنجیره مارکف بشکل متعارفش نیست. با توجه به اینکه متغیرهای X, Y, Z هرکدام با توجه به تمامی متغیرهای قبلی ساخته شده اند، میان دو به دو آنها ارتباط وجود دارد و زنجیره مارکفی میان آنها برقرار نیست. پس آنها را با همدیگر در یک دسته قرار میدهیم. سپس مطابق شکل ۱ متغیر های X و Z را جدا کرده و بقیه متغیرها را از روی آنها میسازیم. در اینجا میبینیم که با حذف Z, X متغیر W از (X, Y, Z) جدا میشود. اما با حذف Z, W متغیر U از Y جدا نمیشود. زیرا زمانی که W را در نظر میگیریم U به X از طریق دسته ای که بدور U و X کشیده شده وصل شده و از طریق آن به Y وصل میشود. با حذف Z متغیر U از (X, Y, Z) جدا میشود، اما متغیر W از (X, Y, Z) جدا نمیشود.

تمرین ۵ یک توزیع مشترک $p(x, y, z, u, w)$ سازگار با روابط مارکف مثال ۴ بسازید بطوریکه $I(U; Y|ZW) = 1$.

۱ نوشتن زنجیره های مارکف برای یک کد

۱.۱ انتقال اطلاعات روی کانال نقطه به نقطه

مساله انتقال اطلاعات روی کانال نقطه به نقطه را در نظر بگیرید. در اثبات وارون از خاصیت بدون حافظه بودن کانال و زنجیره مارکف $MX^nY_{1:i-1} - X_i - Y_i$ استفاده میشود. جهت نوشتن این زنجیره و زنجیره های مشابه میتوان اینگونه



شکل ۲: زنجیره های مارکف برای یک کد انتقال اطلاعات روی کانال نقطه به نقطه

عمل کرد. ما با استفاده از پیام M متغیر X_1، X_n ,$X_2, \cdots$ را میسازیم. سپس با استفاده از X_i متغیر Y_i را میسازیم. پس با استفاده از این ترتیب ساختن، توزیع مشترک این متغیرها را میتوان به این شکل نوشت:

$$p(m, x_1, \cdots, x_n, y_1, \cdots, y_n) = p(m)p(x_1|m)p(x_2|m) \cdots p(x_n|m)p(y_1|x_1)p(y_2|x_2) \cdots p(y_n|x_n).$$

این موضوع در شکل ۲ نشان داده شده است. نحوه ساختن این زنجیره همانند نحوه ساخته شدن این متغیرها از متغیرهای قبلی است. توجه کنید که اگر X_i را از زنجیره ببریم، Y_i از بقیه متغیرها جدا میشود. پس

$$MX_{1:i-1}X_{i+1:n}Y_{1:i-1}Y_{i+1:n} - X_i - Y_i$$

برقرار است. برخی اوقات برای سادگی از نماد $X_{\sim i}$ برای نشان دادن $X_{1:i-1}X_{i+1:n}$ استفاده میکنیم به این معنی که مولفه i ام را به بیرون انداخته ایم. همچنین میبینیم که اگر X_1 و X_2 را ببریم، آنوقت Y_1 و Y_2 از بقیه متغیرها جدا میشوند. پس رابطه زیر نیز برقرار است:

$$MX_{3:n}Y_{3:n} - X_{1:2} - Y_{1:2}. \tag{۱}$$

اگر متغیر Y_3 را هم علاوه بر X_1 و X_2 میبریدیم باز هم Y_1 و Y_2 از بقیه متغیرها جدا میشدند. پس رابطه زیر هم برقرار است.

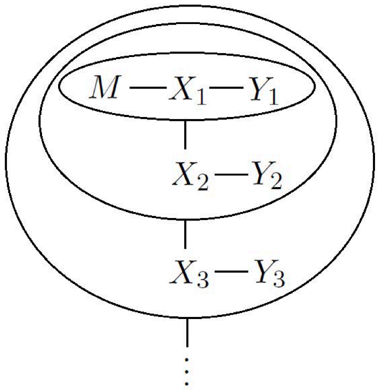
$$MX_{3:n}Y_{4:n} - X_{1:2}Y_3 - Y_{1:2}. \tag{۲}$$

تحقیق کنید که از رابطه (۱) میتوان رابطه (۲) را نتیجه گرفت.

۲.۱ انتقال اطلاعات روی کانال نقطه به نقطه با بازخورد از خروجی

مساله انتقال اطلاعات روی کانال نقطه به نقطه با بازخورد از خروجی را در نظر بگیرید. در اینجا با استفاده از پیام M متغیر X_1 را میسازیم. سپس با استفاده از X_1 متغیر Y_1 را میسازیم. متغیر Y_1 به فرستنده بازخورد داده میشود و فرستنده ورودی X_2 را با توجه به MX_1Y_1 میسازد. سپس با استفاده از X_2 متغیر Y_2 را میسازیم و الی آخر. پس با استفاده از این ترتیب ساختن، توزیع مشترک این متغیرها را میتوان به این شکل نوشت:

$$\begin{aligned} p(m, x_1, \cdots, x_n, y_1, \cdots, y_n) =& p(m)p(x_1|m)p(y_1|x_1)p(x_2|mx_1y_1)p(y_2|x_2) \cdots \\ & p(x_n|mx_{1:n-1}y_{1:n-1})p(y_n|x_n). \end{aligned}$$



شکل ۳: زنجیره های مارکف برای یک کد انتقال اطلاعات روی کانال نقطه به نقطه با بازخورد از خروجی

این موضوع در شکل ۳ نشان داده شده است. نحوه ساختن این زنجیره همانند نحوه ساخته شدن این متغیرها از متغیرهای قبلی است. توجه کنید که اگر X_2 را از زنجیره ببریم، Y_2 از متغیرهای X_1, Y_1, M جدا میشود. پس

$$MX_1Y_1 - X_2 - Y_2$$

برقرار است. اما از طرف دیگر چون X_3 با توجه به تمامی متغیرهای $MX_1Y_1X_2Y_2$ ساخته شده و به همه آنها وصل است، پس بریدن X_2 آن را از Y_2 جدا نمیکند، زیرا X_3 مستقیما به Y_2 وصل است. پس رابطه مارکف

$$X_3 - X_2 - Y_2$$

برقرار نیست. در حالت کلی بریدن X_i، Y_i را از متغیر های قبلی جدا میکند، و نه از متغیرهای بعدی.