

تحلیل مؤلفه‌های اصلی (قسمت دوم)

Principal Component Analysis (Part two)

سجاد طلایی

Talaei.s@arc-ordc.ir

کارشناس ارشد اصلاح نباتات، مرکز تحقیقات کاربردی و تولید بذر، شرکت توسعه کشت دانه‌های روغنی

همچنین می‌توان از داده‌های بیان ژن در بافت‌های مختلف گیاه، پیدا کردن ترکیب‌پذیری خصوصی و عمومی، داده‌های دای آلل، آزمون جدایی‌های تریکودرما بر روی ژنوتیپ‌های مختلف کلزا، بررسی جدایی‌های مختلف زنگ گندم بر روی ژنوتیپ‌های مختلف گندم، مکان‌یابی ژن‌های کنترل‌کننده صفات کمی، تفسیر و نتیجه‌گیری در مورد انتخاب بهترین صفات برای هر ژنوتیپ، بررسی ترکیب‌های تیماری در طرح‌های فاکتوریل نیز استفاده نمود.

ستون‌های ماتریس LEFT بردارهای ویژه چپ و L (ابعاد ماتریس A) که حاوی مقادیر ویژه می‌باشد و ماتریس قطری است و ستون‌های ماتریس RIGHT ردیف‌های آن بردارهای ویژه راست می‌باشند.

در شماره قبل مقدمات محاسبه و بررسی تحلیل مؤلفه‌های اصلی عنوان شد در ادامه می‌بایست به اطلاعات حاصل از تجزیه نرم‌افزار (SAS) پرداخته شود. در خروجی نرم‌افزار مقادیری تحت عنوان LEFT و RIGHT مشاهده می‌شود. این مقادیر بردارهای ویژه (Eigen vector) می‌باشند. در واقع اول مقادیر ویژه محاسبه شده و در فرمول جایگزاری گردیده و بر اساس آنها بردارهای ویژه به دست آمده است. تعداد بردارهای ویژه برابر با تعداد سطر و ستون ماتریس می‌باشند. از این بردارها جهت رسم بای‌پلات استفاده می‌گردد. در برنامه نوشته شده بجای right, left L می‌توان از هر حرفی دیگری استفاده کرد و قانون خاصی ندارد.

در داده‌های حاصل از تحقیقات کشاورزی می‌توان به جای ماتریس سطر - ستون داده‌های ماتریس a از داده‌های دوطرفه استفاده کرد. برای مثال می‌توان از ماتریس داده‌های ژنوتیپ - محیط، سوبه‌های پاتوزن - ژنوتیپ گیاه، ژنوتیپ - صفت و هرگونه ماتریس داده‌های دوطرفه استفاده نمود.

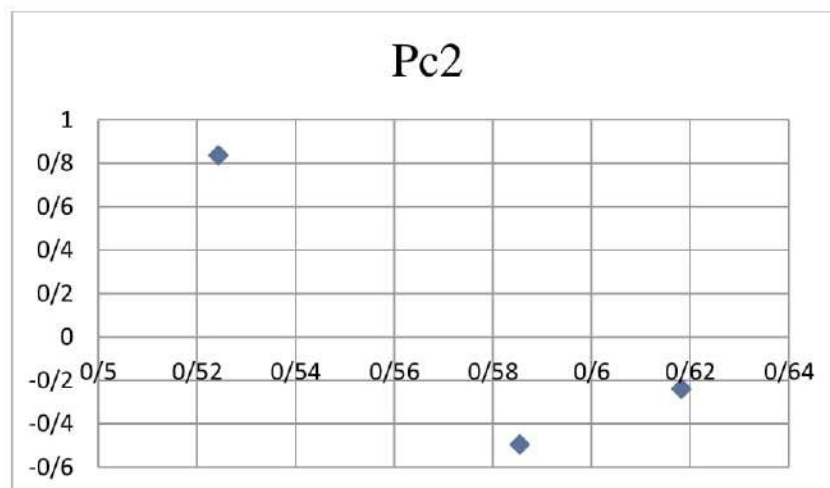
LEFT			RIGHT		
Pc1	Pc2	Pc3	Pc1	Pc2	Pc3
0.5855209	-0.495494	-0.641601	0.9586775	-0.281143	0.043547
0.6182632	-0.238958	0.7487654	0.0678196	0.3744968	0.9247447
0.5243242	0.8350958	-0.166431	0.2762933	0.8835785	-0.378089

در واقع ماتریس A به صورت زیر تفکیک شده است.

$$A = \text{LEFT} \times L \times \text{RIGHT}$$

	LEFT			L			RIGHT			
	Pc1	Pc2	Pc3				Pc1	Pc2	Pc3	
	0.5855209	-0.495494	-0.641601	184.29	0	0	0.9586775	-0.281143	0.043547	
A=	0.6182632	-0.238958	0.7487654	×	0	11.22	×	0.0678196	0.3744968	0.9247447
	0.5243242	0.8350958	-0.166431	0	0	0.40	0.2762933	0.8835785	-0.378089	

بای پلات برای بردارهای ویژه LEFT به صورت زیر حاصل می‌شود.



برخی از مدل‌های آماری پرکاربرد در تجزیه تحلیل داده‌های زیستی مانند SHMM, AMMI و SREG یا همان GGEbiplot بر پایه تجزیه به مقادیر منفرد هستند که در نسخه‌های بعدی به توضیحات بیشتر در خصوص این مدل‌ها پرداخته می‌شود.

منابع:

فرشادفر، ع (۱۳۸۹). اصول و روش‌های تجزیه و تحلیل‌های آماری چند متغیره. انتشارات دانشگاه رازی کرمانشاه. ۷۳۴ص.

Forkman, J. and Piepho, H.P. (2014). Parametric bootstrap methods for testing multiplicative terms in GGE and AMMI models. *Biometrics*. 70(3), 639-647.

Littell, R.C., Milliken, G.A., Stroup, W., Wolfinger, R.D. and Oliver, S. (2006). *SAS for Mixed Models*. SAS Institute. 828 P.