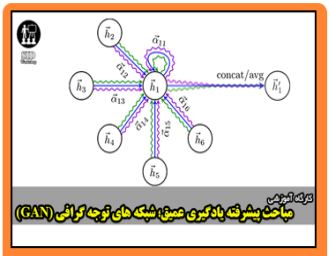
کارگاه های آموزشی مرکز اطلاعات علمی جهاد دانشگاهی

کارگاه آنلاین مکالمه روزمره انگلیسی

کارگاه آنلاین آموزش استفاده از وب آو ساینس



مباحث پیشرفته یادگیری عمیق؛
شبکه های توجه گرافی
(Graph Attention Networks)

چکیده مبسوط پوسترهای ۴۴مین کنفرانس سالانه ریاضی ایران
۵ الی ۸ شهریور ۹۲، دانشگاه فردوسی مشهد، ایران.

تشخیص مشاهدات موثر در برآوردگر لیو و برآوردگر لیو تحت محدودیت تصادفی خطی

فروغ حاجی باقری فروشانی [۱] * و عبدالرحمن راسخ [۲]

گروه آمار، دانشکده علوم ریاضی و کامپیوتر، دانشگاه شهید چمران اهواز
f-hajibagheri@mscstu.scu.ac.ir[۱]
rasekh_a@scu.ac.ir[۲]

چکیده. در آنالیز رگرسیون خطی، تشخیص مشاهدات غیرعادی یک قدم اساسی در جهت فرآیند ساخت مدل است بطوریکه این مشاهدات، تاثیر ناروای زیادی روی آنالیز کمترین مربعات دارد. معیارهای مختلف تشخیص مشاهدات موثر بر پایه استدلال های متفاوت است که روی جنبه های متفاوتی از نتایج رگرسیون های مختلف طراحی شده است. حضور مشاهدات موثر در داده ها زمانی که هم خطی نیز حضور داشته باشد، بسیار پیچیده می شود. در این مقاله، نشان می دهیم هنگامی که از برآوردگرهای لیو و لیو تحت محدودیت تصادفی خطی جهت کاهش اثر هم خطی استفاده می شود تاثیر بعضی مشاهدات را می توان تعمیم داد. یک مثال عددی برای نشان دادن یافته های نظری ارائه گردیده است.

۱. پیش‌گفتار

حضور هم خطی در متغیرهای پیشگو اثری جدی بر برآورد پارامترها و پیش بینی می گذارد. بر این اساس، برآوردگرهای آمیخته و ریج جهت کاهش این اثر پیشنهاد شده اند. علاوه بر هم خطی حضور مشاهدات موثر در داده های مشاهده شده اثر بسیار جدی بر برآوردگرها دارد"در [۱، ۲] ببینید ". هدف اصلی آنالیز تاثیر اندازه گیری تغییرات بوجود آمده در جنبه های مختلف آنالیز، زمانی است که مجموعه ای از داده ها مزاحم (آشفته) باشند. یک طرح قابل توجه از تشخیص داده های مزاحم، حذف موردی است. این طرح در این مقاله مورد استفاده قرارگرفته است.
مدل رگرسیون خطی $y = Z\gamma + \varepsilon$ را در نظر بگیرید که y یک بردار $n \times ۱$ از پاسخ ها، Z یک ماتریس معلوم $n \times p$ از پیشگوها، γ یک بردار $p \times ۱$ از پارامترهای نامعلوم و ε یک بردار $n \times ۱$ از خطاها با $E(\varepsilon) = ۰$ و $Var(\varepsilon) = \sigma^۲ I_n$ باشد. در این صورت برآوردگر کمترین مربعات معمولی برای γ برابر است با $\widehat{\gamma} = (Z'Z)^{-۱}Z'y$. برآورد $\sigma^۲$ برابر است با

2010 Mathematics Subject Classification. Primary 62J07; Secondary 62J20.
واژگان کلیدی. برآوردگر لیو، برآوردگر لیو تحت محدودیت تصادفی خطی، مشاهدات موثر، مباحث تشخیصی، هم خطی .
* سخنران

$s^۲ = e'e/(n-p)$، که $e = y - Z\widehat{\gamma}$ بردار باقی مانده هاست.
مدل رگرسیون خطی $y = X\beta + \varepsilon$ را فرم متعارف مدل $y = Z\gamma + \varepsilon$ گوییم به این صورت که $X = ZT$، $\beta = T'\gamma$ و T ماتریس متعامدی است که ستون هایش بردارهای ویژه $Z'Z$ را تشکیل می دهد. بنابراین $X'X = T'Z'ZT = \Lambda = diag(\lambda_۱, ..., \lambda_p)$، که $\lambda_۱ \geq ... \geq \lambda_p > ۰$ مقادیر ویژه $Z'Z$ هستند. در این صورت برآوردگر کمترین مربعات را می توان به این صورت نوشت $\widehat{\beta} = \Lambda^{-۱}X'y$ بطوریکه $\widehat{\gamma} = T\widehat{\beta}$.
موقعیت یا اثر $i-$امین مشاهده با h_i اندازه گیری می شود، که $i-$امین درایه قطری از ماتریس برازش $H = X\Lambda^{-۱}X'$ است.
برآوردگر لیو جهت کاهش اثر هم خطی بسیار مفید است. "در [۳] ببینید ". هدف اصلی این مقاله تعمیم مباحث تشخیصی به برآوردگر لیو و برآوردگر لیو تحت محدودیت تصادفی خطی است.

تعریف ۱.۱. *برآوردگر لیو $\widehat{\beta}_d$ توسط لیو "در [۳] ببینید "به صورت زیر تعریف شده است: $\widehat{\beta}_d = (\Lambda + I)^{-۱}(X'y + d\widehat{\beta})$ که I یک ماتریس همانی است، $\widehat{\beta}$ برآوردگر کمترین مربعات متعارف و $۰ < d < ۱$ پارامتر اریبی برآوردگر لیو است. این برآوردگر ترکیبی از برآوردگرهای رگرسیونی ریج و استاین است.*

تعریف ۲.۱. *برآوردگر لیو تحت محدودیت تصادفی خطی $\widehat{\beta}_{srd}$ توسط هابرت و ویجیکون"در [۴] ببینید" به صورت زیر معرفی شد :*

$$\widehat{\beta}_{srd} = (\Lambda + I)^{-۱}(\Lambda + dI)\widehat{\beta}_m$$

که $\widehat{\beta}_m = \widehat{\beta} + \Lambda^{-۱}R'(I + R\Lambda^{-۱}R')^{-۱}(r - R\widehat{\beta})$ برآوردگر آمیخته ای است که از اضافه کردن محدودیت تصادفی خطی $r = R\gamma + \phi$ به ماتریس مشاهدات نمونه با شروط $E(\phi) = ۰$، $Var(\phi) = \sigma^۲ I$ و $E(\varepsilon\phi') = ۰$ به دست می آید.

تعریف ۳.۱. *عمومی ترین معیاری که می تواند اثر یک مشاهده را اندازه گیری کند تفاوت در مقادیر برازش استاندارد شده است "در [۱] ببینید ". $i-$امین مورد آن به این صورت محاسبه می شود*

$$DFFITS(i) = x_i[\widehat{\beta} - \widehat{\beta}(i)]/SE(x_i\widehat{\beta})$$

که $\widehat{\beta}(i)$ برآوردگر کمترین مربعات از $\widehat{\beta}$ است که $i-$امین مورد آن حذف شده است و $SE(x_i\widehat{\beta})$ برآوردگر خطای استاندارد مقادیر برازش شده است. $DFFITS$ تغییر استاندارد شده در مقادیر برازش شده در اثر حذف یک مشاهده است.

تعریف ۴.۱. *دیگر معیار مفید تاثیر، فاصله کوک می باشد که $i-$امین مورد آن به این صورت محاسبه می شود*

$$D_i = \frac{۱}{ps^۲}[\widehat{\beta} - \widehat{\beta}(i)]'(X'X)[\widehat{\beta} - \widehat{\beta}(i)]$$

که D_i معیار تغییر در همه مقادیر برازش شده است، هنگامی که یک مشاهده حذف شده باشد. نقاطی با مقادیر بزرگ D_i از نظر تاثیر روی برآوردگر کمترین مربعات $\widehat{\beta}$ قابل توجه هستند.

۲. دست‌آوردهای پژوهش

ابتدا معیارهای آنالیز تاثیر را برای برآوردگر لیو و سپس برآوردگر لیو تحت محدودیت تصادفی خطی به دست می آوریم و سپس یافته ها را با یک مثال عددی نشان می دهیم.
بردار مقادیر برازش شده برآوردگر لیو عبارت است از

$$\begin{aligned}\widehat{y}_d &= X\widehat{\beta}_d = X(\Lambda+I)^{-1}(X'y+d\widehat{\beta}) \\ &= X(\Lambda+I)^{-1}(\Lambda+dI)\Lambda^{-1}X'y = H_d y\end{aligned}$$

که $H_d = X(\Lambda+I)^{-1}(\Lambda+dI)\Lambda^{-1}X'$ ماتریس برازش برآوردگر لیو است و نقشی شبیه به ماتریس برازش در برآوردگر کمترین مربعات بازی می کند. توجه به این نکته مهم است که H_d یک ماتریس تصویرگر نیست زیرا خودتوان نیست و H_d ماتریس شبه تصویرگر نامیده می شود "در [۲] ببینید". $\widehat{y}_{di} = \sum_{j=1}^{n} h_{dij}y_j$ $i-$امین مقدار برازش شده ی برآوردگر لیو برابر است
در این صورت بردار باقی مانده های برآوردگر لیو برابر است با $\widehat{\varepsilon}_d = y - \widehat{y}_d = (I - H_d)y$
.

$DFFITS$ را برای برآوردگر لیو می توان به صورت زیر نوشت

$$DFFITS_d(i) = x_i[\widehat{\beta}d - \widehat{\beta}_d(i)]/SE(x_i\widehat{\beta}_d)$$

که $\widehat{\beta}_d(i)$ برآوردگر لیو پس از حذف $i-$امین مشاهده می باشد و مخرج کسر برابر است با
$SE(x_i\widehat{\beta}_d) = s[x_i(\Lambda+I)^{-1}(\Lambda+dI)\Lambda^{-1}(\Lambda+dI)'(\Lambda+I)^{-1}x_i']^{1/2}$.
چون میانگین مربع خطا تابعی از مقادیر برازش شده و پاسخ است و به تک تک مقادیر ویژه ی $Z'Z$ وابسته نیست پس تحت تاثیر هم خطی نیست. به همین دلیل s برآوردگر کمترین مربعات σ هنوز به عنوان معیار مقیاس مورد استفاده قرار می گیرد.
دو گونه مختلف از فاصله کوک را می توان به صورت زیر به دست آورد

$$D_{di} = \frac{1}{ps^2}[\widehat{\beta}_d - \widehat{\beta}_d(i)]'\Lambda[\widehat{\beta}_d - \widehat{\beta}_d(i)]$$

$$D'_{di} = \frac{1}{ps^2}[\widehat{\beta}_d - \widehat{\beta}_d(i)]'(\Lambda+I)(\Lambda+dI)^{-1}\Lambda(\Lambda+dI)^{-1}(\Lambda+I)'[\widehat{\beta}_d - \widehat{\beta}_d(i)]$$

که D_{di} یک تعمیم مستقیم از فاصله کوک تعریف ۴.۱ است و D'_{di} بر پایه این حقیقت است که

$$Var(\widehat{\beta}_d) = \sigma^2(\Lambda+I)^{-1}(\Lambda+dI)\Lambda^{-1}(\Lambda+dI)'(\Lambda+I)^{-1}$$

بردار مقادیر برازش شده برای $\widehat{\beta}_{srd}$ برابر است با $\widehat{y}_{srd} = X\widehat{\beta}_{srd}$ در این صورت بردار باقی مانده های این برآوردگر برابر است با $\widehat{\varepsilon}_{srd} = y - \widehat{y}_{srd}$.

$DFFITS$ را برای این برآوردگر می توان به صورت زیر نوشت

$$DFFITS_{srd}(i) = [\widehat{\beta}_{srd} - \widehat{\beta}_{srd}(i)]/SE(x_i\widehat{\beta}_{srd})$$

که

$$SE(x_i\widehat{\beta}_{srd}) = s[x_i(\Lambda+I)^{-1}(\Lambda+dI)(\Lambda+R'R)^{-1}(\Lambda+dI)'(\Lambda+I)^{-1}x_i']^{1/2}$$

و $Var(\widehat{\beta}_m) = \sigma^2(\Lambda + R'R)^{-1}$.

فواصل کوک برآوردگر لیو تحت محدودیت را نیز می توان به صورت زیر به دست آورد

$$D_{srdi} = \frac{1}{ps^2}[\widehat{\beta}_{srd} - \widehat{\beta}_{srdi}(i)]'\Lambda[\widehat{\beta}_{srd} - \widehat{\beta}_{srdi}(i)]$$

$$D'_{srdi} = \frac{1}{ps^2}[\widehat{\beta}_{srd} - \widehat{\beta}_{srdi}(i)]'(\Lambda + I)(\Lambda + dI)^{-1} \times(\Lambda + R'R)(\Lambda + dI)^{-1}(\Lambda + I)[\widehat{\beta}_{srd} - \widehat{\beta}_{srdi}(i)]$$

در این بخش برای نشان دادن نتایج نظری از داده های سیمان پورتلند (هالد) "در [۵] ببینید "که از یک تحقیق تجربی از حرارت تکامل یافته در طول سخت شدن سیمان و وابستگی این گرما روی درصد چهار ترکیب با ۱۳مشاهده استفاده شده است. شرط عددی هم خطی با عرض از مبدا ۶۰۵۶ است که با اعمال محدودیت لیو این مقدار به ۲۱۱ کاهش می یابد. محدودیت تصادفی ارائه شده برای این داده ها $\beta = 0$(۱،۰ ،-۱،۰ ،۱،۰) است. با توجه به مقدار برآوردها "در [۵] ببینید " معیارهای تشخیص، موثرترین مشاهدات را نشان می دهد.

جدول ۱: چهارتا از موثرترین مشاهدات در داده های هالد(به ترتیب از چپ)

$DFFITS$	D	D'
$\widehat{\beta}$ ۸،۳،۱۱،۶	۸،۳،۱۱،۱۳	۸،۳،۱۱،۱۳
$\widehat{\beta}_d$ ۸،۱۳،۶،۱۱	۸،۱۳،۶،۱۱	۸،۱۳،۱۱،۶
$\widehat{\beta}_{srd}$ ۸،۳،۱۱،۶	۸،۳،۱۱،۶	۸،۳،۶،۱۱

سطرهای این جدول به ترتیب موثرترین مشاهدات را بر اساس بزرگی مقادیر روش تشخیصی مورد استفاده، نسبت به بقیه مشاهدات؛ در به کار گیری برآوردگرهای کمترین مربعات، لیو و لیو تحت محدودیت تصادفی نشان می دهد که تفاوت نقاط موثر را با اختلاف کمی مشاهده می فرمایید اما در هر سه روش برآورد و هر سه روش تشخیص، مشاهده ی ۸ام دارای بیشترین تاثیر است. گرچه هیچ نقطه برشی برای اثرات تشخیصی ارائه شده معرفی نشد اما ارائه نمودار که روشی بصری و مرسوم در افشای موارد تاثیرگذار است، می تواند مفید واقع شود.

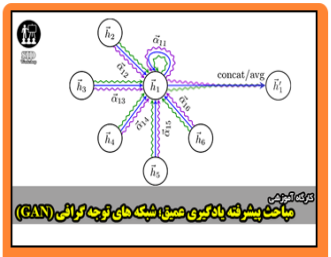
مراجع

1. D.A. Belsley, E. Kuh and R.E. Welsch, *Regression diagnostics: identify influence data and source of collinearity*, Wiley, New York, 1980.
2. E. Walker and J.B. Birch, *Influence measure in ridge regression*, Technometrics. 30 (1988), 221–227.
3. K. Liu, *A new class of biased estimate in linear regression*, Commun. Statist. Theo. Meth. 22 (1993), 393–402.
4. M.H. Hubert and P. Wijekoon, *Improvement of the Liu estimator in linear regression model*, Stat. Pape. 47 (2006), 471–479.
5. S. Kaciranlar, S. Sakallioglu, F. Akdeniz, G.P.H. Styan and H.J. Werner, *A new biased estimator in linear regression and a detailed analysis of the widely analyzed dataset on Portland Cement*, Sankhya: Ind. J. of Stat. 61B (1999), 443–459.

کارگاه های آموزشی مرکز اطلاعات علمی جهاد دانشگاهی

کارگاه آنلاین مکالمه روزمره انگلیسی

کارگاه آنلاین آموزش استفاده از وب آو ساینس



مباحث پیشرفته یادگیری عمیق؛ شبکه های توجه گرافی (Graph Attention Networks)