بسم الله الرحمن الرحیم

بهار

دانشگاه یزد

دانشکده ریاضی

گروه آمار

پایان‌نامه

برای دریافت درجه کارشناسی ارشد
آمار ریاضی

استنباط بیزی در مدل تحلیل واریانس یکطرفه

استاد راهنما: دکترعلی دولتی

استاد مشاور: دکتر عیسی محمودی

پژوهش و نگارش: جلال زین الدینی

بهمن ماه ۱۳۹۱

چکیده

یکی از ابزارهای اساسی در تحلیل داده‌های حاصل از طراحی آزمایشات، تحلیل واریانس است. در تحلیل واریانس یک‌طرفه، هدف مقایسه میانگین چند جامعه مستقل است. در روش کلاسیک، آماره آزمون مورد استفاده، دارای توزیع F مرکزی و تحت فرض مقابل دارای توزیع F با یک پارامتر نامرکزی است. در این پایان‌نامه ابتدا مسأله آزمون فرض تساوی میانگین‌ها که فرضیه‌ای چند پارامتری است، به مسأله آزمون فرض صفر بودن پارامتر نامرکزی توزیع آماره آزمون که فرضیه‌ای تک پارامتری و ساده‌تر است، تبدیل می‌شود. سپس با در نظر گرفتن توزیع پیشین برای پارامترها و انعکاس آن در پارامتر نامرکزی، به تحلیل واریانس یک‌طرفه و استنباط‌های مربوط به آن با رویکرد آمار بیز پرداخته می‌شود.

فهرست مطالب

۱ مفاهیم و تعاریف اولیه ۵

۱.۱ مفاهیمی از آنالیز ریاضی ۶

۲.۱ مفاهیمی از احتمال ۸

۳.۱ مفاهیمی از جبر خطی ۹

۱.۳.۱ مشتق تابع حقیقی مقدار نسبت به ماتریس ۱۱

۲.۳.۱ مشتق بردار نسبت به بردار ۱۲

۳.۳.۱ ژاکوبی تبدیل‌های ماتریس ۱۲

۴.۱ استنباط آماری ۱۴

۱.۴.۱ استنباط کلاسیک ۱۴

۲.۴.۱ استنباط بیزی ۱۵

۲ برآورد در مدل خطی کلاسیک و بیزی ۲۷

۱.۲ برآوردهای کمترین مربعات خطا ۲۸

۲.۲ ویژگی‌های برآوردگرهای کمترین مربعات خطا ۲۹

۳.۲ برآوردهای بیز برای مدل خطی استاندارد ۳۰
۴.۲ برآوردگرهای بیز ۳۳
۵.۲ تحلیل واریانس یک‌طرفه ۳۴
۶.۲ برآورد پارامترها ۳۷
۱.۶.۲ برآوردهای کلاسیک مدل تحلیل واریانس یک‌طرفه استاندارد ۳۷
۲.۶.۲ برآوردهای بیز برای مدل تحلیل واریانس استاندارد ۳۹

۳ استنباط بیزی در مدل تحلیل واریانس یک‌طرفه **۴۱**

۱.۳ فاکتور بیز ۴۲
۱.۱.۳ نمونه آموزشی ۴۳
۲.۳ فاکتور بیز ذاتی ۴۵
۱.۲.۳ مزایای کاربرد IBF و توصیه‌های برگر و پریچی ۴۸
۲.۲.۳ فاکتور بیز کسری ۴۹
۳.۳ رویکرد درستنمایی حاشیه‌ای در تحلیل واریانس ۴۹
۴.۳ رویکرد درستنمایی تلفیقی در تحلیل واریانس ۶۲
۱.۴.۳ پارامترسازی مجدد از $(\mu, \sigma^۲)$ ۶۵
۲.۴.۳ توزیع‌های پیشین مرجع ۷۰
۳.۴.۳ روش درستنمایی تلفیقی ۷۴
۵.۳ انتخاب بین توزیع های پیشین بدون اطلاع ۷۸
۱.۵.۳ توزیع پیشین تطابقی مرتبه اول برای τ ۷۸

۴ تحلیل داده‌ها و شبیه‌سازی ۸۱
۱.۴ مقدمه ۸۲
۲.۴ مقایسه توزیع‌های پیشین $\pi_r(\tau)$ و $\pi_d(\tau)$ برای نمونه‌های کوچک ۸۲
۱.۲.۴ تئوری زنجیر مارکوف ۸۳
۲.۲.۴ الگوریتم متروپولیس-هستینگ ۸۵
۳.۴ تحلیل داده‌ها با استفاده از فاکتور بیز ۹۰

الف برنامه‌های کامپیوتری ۹۵

ب واژه‌نامه فارسی به انگلیسی ۱۰۵

پ واژه نامه انگلیسی به فارسی ۱۱۱

مراجع ۱۱۷

لیست تصاویر

۱.۳ درستنمایی نسبی حاشیه‌ای و کامل ۵۶

۲.۳ درستنمایی حاشیه‌ای نسبی و کامل ۵۷

۳.۳ رابطه بین F_{obs} و $\overline{M}(۰)$ و $\overline{L}(۰)$ ۵۸

۴.۳ رابطه بین F_{obs} و $\hat{\lambda}$ و $\widetilde{\lambda}$ و محدوده‌های فواصل درستنمایی ۵۸

۵.۳ نقاط $(\alpha_{obs}, \overline{L}(۰))$ برای $d_۱ = ۲$، $d_۲ = ۳, ۸, \infty$ و $d_۱ = ۷$، $d_۲ = ۸, ۱۶, \infty$ ۵۹

۶.۳ تابع درستنمایی حاشیه‌ای مربوط به مقدار $\widetilde{\tau}$ برای $d_۲ = ۴, ۸, ۱۶, ۳۲, ۶۴$ . . ۶۲

۷.۳ رابطه بین $d_۲$ و $\hat{\lambda}$، $\widetilde{\lambda}$ و $\hat{\tau}$، τ وقتی $\tau = ۰/۲۸$ ۶۳

فهرست نمادها

برآوردگر درستنمایی ماکزیمم .... MLE

تابع چگالی احتمال .... f

تابع توزیع .... F

توزیع پیشین .... $\pi(\theta)$

توزیع پسین .... $\pi(\theta \mid x)$

گوی باز به مرکز p و شعاع r .... $B_r(p)$

بستار مجموعه A .... $\overline{A}$

ترانهاده مجموعه A .... A^t

متمم مجموعه A .... A^c

درجه آزادی توزیع t .... ν

فضای نمونه .... Ω

میدان اعداد حقیقی .... $\mathbb{R}$

تابع درستنمایی .... $L(.)$

لگاریتم تابع درستنمایی .... $l(.)$

برآوردگر درستنمایی ماکزیمم θ .... $\hat{\theta}$

مقدار اولیه برای پارامتر θ .... θ^0

مشتق جزئی .... ∂

توزیع t، p متغیره .... t_p

فاصله بین مشاهدات p و q .... d(p, q)

معیار اطلاع بیزی .... BIC

معیار اطلاع آکائیک .... AIC

رتبه ماتریس A .... $Rank(A)$

معکوس ماتریس A .... A^{-1}

ژاکوبی تبدیل U به V .... $j(U \to V)$

اطلاع فیشر .... $I_x(\theta)$

روشهای زنجیر مارکوف مونت کارلو .... $MCMC$

تناسب .... α

مخاطره برآورد δدر برآورد پارامتر θ .... $R(\theta, \delta)$

برآوردگر ماکزیمم پسین .... MAP

بردار پارامترهای مجهول .... $\underline{\beta}$

بردار مؤلفه‌های خطا .... $\underline{\epsilon}$

فاکتور بیز ذاتی .... IBF

فاکتور بیز ذاتی با اصلاح حسابی .... $AIBF$

فاکتور بیز ذاتی با اصلاح هندسی .... $GIBF$

فاکتور بیز کسری .... FBF

تابع درستنمایی حاشیه‌ای نسبی λ .... $\overline{L}(\lambda)$

تابع درستنمایی کامل λ .... $M(\lambda)$

تابع درستنمایی کامل نسبی λ .... $\overline{M}(\lambda)$

مشتق جزئی مرتبه دوم .... $\partial^{۲}$

تابع پوشش هامر .... $(q)_n$

بردار گرادیان تابع g .... $\bigtriangledown_g$

بردار پارامترهای مزاحم .... $\mathfrak{S}$

پیشگفتار

امروزه بسیاری از شرکت‌های مدرن علمی با این سوال که چه مدلی را باید انتخاب کنند و طبق آن عمل نمایند روبه‌رو هستند. آزمایشگر یا گردآورنده داده‌ها، اغلب داده‌ها را از جنبه‌های مختلف ارزیابی می‌کند تا تأثیر متغیرها را بر روی پیامدهای مورد نظر خود بیابد. کدام اندازه‌ها روی نتیجه آزمایش تأثیر بیشتری دارند؟ کدام یک از آنها اهمیت چندانی ندارند؟ آیا اثرات متقابلی بین متغیرها وجود دارد که باید در نظر گرفته شود؟ به عنوان مثال فرض می‌شود که قرار است تأثیر یک مداخله پزشکی در تعداد دو نوع سلول بررسی شود. بدین منظور تعداد هر دو نوع سلول، قبل و بعد از مداخله برای N فرد به دست آمده است. سپس برای هر دو نوع سلول تعداد ثبت شده قبل از مداخله از تعداد بعد از مداخله کسر گردیده است. نتایج حاصل برای دو نوع سلول با (y_{i1}, y_{i2})، $(i = 1, ..., N)$ نشان داده می شود که y_{i1} و y_{i2} به ترتیب نشانگر اختلاف‌ها در تعداد سلول نوع یک و نوع دو هستند. وقتی که توزیع تعداد اختلاف‌ها یعنی توزیع (y_{i1}, y_{i2}) دارای یک شکل بیضوی است، داده‌ها را می‌توان با استفاده از یک مدل نرمال دو متغیره که در زیر داده شده است تقریب زد:

$$\begin{bmatrix} y_{i1} \\ y_{i2} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$$

که $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ میانگین‌های تعداد اختلاف‌ها و ماتریس 2×2 نشان داده شده ماتریس کوواریانس خطا است که با Σ نشان داده می‌شود. انتظارات سه متخصص پزشکی برای این مسأله به صورت زیر است. اولین فرد ادعا می‌کند که مداخله هیچ تأثیری روی تعداد سلول‌ها ندارد. متخصص دوم اظهار می‌دارد که مداخله روی تعداد سلول‌ها دارای تأثیر مثبت است و سومین متخصص معتقد است که مداخله در تعداد سلول‌های نوع دوم تأثیر بیشتری داشته است یا به عبارت دقیق‌تر تعداد سلول‌های بعد از مداخله منهای تعداد سلول‌های قبل از مداخله برای سلول نوع دوم بزرگتر از سلول نوع یک است. بر اساس مدل فوق ادعاهای مطرح شده را می‌توان به صورت زیر فرمول‌بندی کرد:

$$\begin{cases} \mathcal{M}_1 : \mu_1 = 0, \mu_2 = 0 \\ \mathcal{M}_2 : \mu_1 > 0, \mu_2 > 0 \\ \mathcal{M}_3 : \mu_2 > \mu_1 \end{cases}$$

ضمناً ماتریس کوواریانس در اینجا برای ساده‌تر شدن نمادگذاری نادیده گرفته شده است. این مثال در واقع به نوعی اهمیت انتخاب مدل در مسائل و همچنین مشکلاتی که در این راستا وجود دارد را خاطر نشان می‌سازد. اولین مسأله آن است که مدل‌ها شامل محدودیت‌های برابری و نابرابری بین پارامترهای مدل هستند. این محدودیت‌ها در واقع ترجمه ریاضی اظهارات اعلام شده توسط پژوهشگران هستند. از طرفی مدل‌های مورد نظر می‌توانند آشیانه‌ای یا غیرآشیانه‌ای باشند و یا حتی ممکن است تا حدودی آشیانه‌ای باشند و همه مدل‌ها در یک مدل $\mathcal{M}_0$ که در نظر گرفته نشده است آشیانه شده باشند. در مثال فوق $\mathcal{M}_0 : (\mu_1, \mu_2)' \in \mathbb{R}^2$. از طرفی در برخی موارد مجموعه مدل‌های تحت بررسی $\mathcal{M}_1, \ldots, \mathcal{M}_T$ ممکن است شامل بیش از دو مدل باشد، به عبارت دیگر T ممکن است بزرگتر از دو باشد. در واقع یکی از اهداف علم آمار تأمین ابزارهایی است که بتواند در پاسخ به این‌گونه سوالات یاریگر محقق باشد. آماردانان نیز به طور طبیعی در پاسخ به سوال انتخاب مدل، شرکت داشته‌اند و تاکنون روش‌های گوناگونی برای پاسخ به این سوال ارائه شده است. در رابطه با مسأله انتخاب مدل، هم آماردانان کلاسیک و هم آماردانان بیزی نقش بسزایی دارند. روش‌هایی مانند آزمون‌های F برای مدل‌های آشیانه‌ای، معیار اطلاع آکائیک (AIC)، روش‌های جست و جوی جامع، قدم به قدم، پس‌رو و پیش‌رو، اعتبارسنجی متقابل، فاکتورهای بیز (جزئی، ذاتی، کسری و پسین)، اندازه اطلاع بیز (BIC) و متوسط مدل بیزی نام برخی از روش‌های محبوب و شناخته شده برای انتخاب مدل است. برخی از این روش‌ها مانند جست و جوی قدم به قدم الگوریتمهایی برای ساختن یک مدل خوب (یا مفید) هستند و برخی از روش‌ها مانند معیار اطلاع آکائیک، معیارهایی برای ارزیابی کیفیت مدل می‌باشند. با چنین غنایی که در انتخاب روش وجود دارد، آماردانان چگونه تصمیم می‌گیرند که چه کاری انجام دهند؟ روشی که نتواند توسط جامعه علمی درک شود و یا به اجرا درآید، پذیرفته نیست. بنابراین به روشی نیاز است که پیاده‌سازی آن دشوار نباشد و نتایج حاصل از آن توسط متخصصان علمی به راحتی قابل تفسیر و قابل شرح به صورت عددی برای کاربران نهایی باشد. از دیدگاه آماری به روشی منسجم و به اندازه کافی کلی نیاز است تا بتواند بسیاری از مسائل را پوشش دهد [۲۵]. در این پایان‌نامه، ابتدا در فصل اول به برخی از تعاریف و مفاهیمی که در فصل‌های آتی مثمر ثمر خواهد بود پرداخته می‌شود. سپس در فصل دوم برآورد مدل خطی به روش کلاسیک و بیزی بیان می‌شود. در فصل سوم نیز استنباط بیزی در مدل تحلیل واریانس یک‌طرفه به

تفصیل مطرح خواهد شد و در نهایت در فصل چهارم با ارائه یک مثال عددی نتایج کلی پایان‌نامه نشان داده می‌شوند. مرجع اصلی این پایان نامه، مرجع [۴۶] می‌باشد.

فصل ۱

مفاهیم و تعاریف اولیه

مقدمه

در این فصل برخی از تعاریف و قضایا که در فصل‌های آتی مورد استفاده قرار می‌گیرند به‌طور خلاصه یادآوری شده است. در بخش ۲.۱ مفاهیمی از آنالیز ریاضی، در بخش ۳.۱ مفاهیمی از احتمال، در بخش ۴.۱ مفاهیمی از جبرخطی و در بخش ۵.۱ مفاهیمی از استنباط آماری (کلاسیک و بیزی) یادآوری می‌شوند.

۱.۱ مفاهیمی از آنالیز ریاضی

در این بخش به بیان تعاریف و قضایایی از آنالیز ریاضی پرداخته می شود و برای جزئیات به [۵۵] ارجاع داده می شود.

تعریف ۱.۱.۱. *فضای متریک عبارتست از زوج (S,D) که S یک مجموعه ناتهی از نقاط و D یک تابع فاصله یا متریک از $S\times S$ به $[0,\infty)$ با ویژگی‌های زیر است:*

الف) $D(p,q)=0$ *اگر و فقط اگر* $p=q$،

ب) $D(p,q)>0$ *اگر* $p\neq q$،

ج) $D(p,q)=D(q,p)$ *به ازای هر* $p,q\in S$،

د) *به ازای هر* $p,q,r\in S$ ؛ $D(p,r)\leq D(p,q)+D(q,r)$.

اگر تابع D ویژگی ۴ را نداشته باشد، آن را یک فاصله می‌نامند. اگر هر چهار ویژگی را داشته باشد به آن یک متر یا متریک می‌گویند.

تعریف ۲.۱.۱. *اگر (S,D) یک فضای متریک باشد و $p\in S$، به ازای $r>0$ مجموعه $B_r(p)=\{x|D(x,p)<r\}$ را یک گوی باز به مرکز p و شعاع r یا یک همسایگی از p می‌نامند.*

تعریف ۳.۱.۱. *مجموعه‌ی* $\bar{A} = A \cup A^t$ *بستار* A *نامیده می شود.*

تعریف ۴.۱.۱. *مجموعه* A *را بسته گویند اگر* $A = \bar{A}$. *مجموعه را باز گویند اگر* A^c *بسته باشد یا بطور معادل اگر برای هر* $p \in A$ *وجود داشته باشد* $r > ۰$ *بطوری‌که* $B_r(p)$ *در* A *قرار بگیرد.*

تعریف ۵.۱.۱. *(تابع فوق هندسی همشارکومر) سری توانی نامتناهی مطلقاً همگرای زیر، تابع فوق هندسی همشارکومر نام دارد.*

$$M(a,b,z) = \sum_{j=0}^{\infty} \frac{(a)_j}{(b)_j} \frac{z^j}{j!}$$

تعریف ۶.۱.۱. *سری فوق هندسی گاوس برای* $|z| < ۱$ *به صورت زیر تعریف می‌شود:*

$$F(a,b;c;z) = \sum_{i=۱}^{\infty} \frac{(a)_n (b)_n}{c_n} \frac{z^n}{n!}$$

که در آن c *مثبت و* $(q)_n$ *نماد تابع پوش‌هامر است که به صورت زیر تعریف می‌گردد:*

$$(q)_n = \begin{cases} ۱ & \text{اگر } n = ۰ \\ q(q+۱)...(q+n-۱) & n > ۰ \end{cases}$$

یک روش استاندارد برای مقایسه‌ی دو دنباله یا دو تابع، استفاده از نمادهای o کوچک و O بزرگ است که در زیر به آن اشاره می‌شود [۵۷].

تعریف ۷.۱.۱. *(نماد* O *بزرگ) اگر* $\{x_n\}$ *و* $\{a_n\}$ *دو دنباله‌ی مختلف باشند، آن‌گاه*

$$x_n = \boldsymbol{O}(a_n),$$

اگر اعداد ثابت C *و* $n_۰$ *وجود داشته باشد به قسمی‌که،*

$$|x_n| \leq C|a_n|,$$

وقتی‌که $n \geq n_۰$.

تعریف ۸.۱.۱. *(نماد $\boldsymbol{o}$ کوچک) اگر $\{x_n\}$ و $\{a_n\}$ دو دنباله‌ی مختلف باشند، آن‌گاه*

$$x_n = \boldsymbol{o}(a_n),$$

اگر

$$\lim_{n\to\infty} \frac{x_n}{a_n} = ۰.$$

برای دو تابع f و g نتایج زیر برقرار است:

۱) $\boldsymbol{O}(f(x)+g(x)) = \boldsymbol{O}(f(x)) + \boldsymbol{O}(g(x))$.

۲) $\boldsymbol{O}(f(x)g(x)) = \boldsymbol{O}(f(x))\boldsymbol{O}(g(x))$.

۳) $\boldsymbol{o}(f(x)g(x)) = \boldsymbol{o}(f(x))\boldsymbol{o}(g(x))$.

۴) اگر $f(x)$ و $g(x)$ زمانی که $x \to a$ هم‌ارز باشند (یعنی $\lim_{x\to a} \frac{f(x)}{g(x)} = ۱$)، آن‌گاه:

$$f(x) = g(x) + \boldsymbol{o}(g(x)), \quad x \to a.$$

۲.۱ مفاهیمی از احتمال

(روش دلتا) اگر $g : \mathbb{R}^p \longrightarrow \mathbb{R}^k$ $(k \leq p)$ به صورتی تعریف شده باشد که در نقطه $\boldsymbol{\mu} \in R^p$ پیوسته باشد و X_n دنباله‌ای از بردارهای pبعدی باشد که

$$\sqrt{n}\left(X_n - \boldsymbol{\mu}\right) \xrightarrow{d} N_p\left(۰, \Sigma\right),$$

در این‌صورت

$$\sqrt{n}\left(g(X_n) - g(\boldsymbol{\mu})\right) \xrightarrow{d} N_k\left(۰, \nabla g(\boldsymbol{\mu})\Sigma\left(\nabla g(\boldsymbol{\mu})\right)^t\right),$$

که در آن ∇g بردار گرادیان تابع g است.

تعریف ۱.۲.۱. *متغیر تصادفی $\boldsymbol{X}$ دارای توزیع نرمال p متغیره با پارامترهای $(\boldsymbol{\mu}, \Sigma)$ به صورت $\boldsymbol{X} \sim \boldsymbol{N}_p(\boldsymbol{\mu}, \Sigma)$ نشان داده می شود و تابع چگالی آن به صورت زیر است:*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{۱}{\sqrt{(۲\pi)^p|\Sigma|}} \exp(-\frac{۱}{۲}(\boldsymbol{x}-\boldsymbol{\mu})^T|\Sigma|^{-۱}(\boldsymbol{x}-\boldsymbol{\mu}))$$

تعریف ۲.۲.۱. *متغیر تصادفی X دارای توزیع t، p متغیره با پارامترهای $(\Sigma, \boldsymbol{\mu}, \nu)$ به صورت $\boldsymbol{X} \sim \boldsymbol{t}_p(\Sigma, \boldsymbol{\mu}, \nu)$ نشان داده می شود و تابع چگالی آن به صورت زیر است:*

$$f_{\boldsymbol{X}}(x_۱, \ldots, x_p) = \frac{\Gamma[(\nu+p)/۲]}{\Gamma(\nu/۲)\nu^{p/۲}\pi^{p/۲}|\boldsymbol{\Sigma}|^{۱/۲}[۱+\frac{۱}{\nu}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-۱}(\boldsymbol{x}-\boldsymbol{\mu})]^{(\nu+p)/۲}}$$

(قضیه‌ی حد مرکزی چند متغیره) اگر $X_۱, X_۲, \ldots, X_n$ یک نمونه‌ی تصادفی n تایی از بردار p بعدی X با بردار میانگین $\boldsymbol{\mu}$ و ماتریس کواریانس Σ باشد، در این‌صورت

$$\sqrt{n}\left(\overline{X}_n - \boldsymbol{\mu}\right) \xrightarrow{d} N_p\left(۰, \Sigma\right),$$

که در آن

$$\overline{X}_n = \left(\frac{1}{n}\sum_{j=1}^{n} X_{1j}, \ldots, \frac{1}{n}\sum_{j=1}^{n} X_{pj}\right)^t$$

بردار میانگین نمونه‌ای است.

۳.۱ مفاهیمی از جبر خطی

در این بخش برخی از مفاهیم مورد نیاز مانند مشتق و ماتریس از جبرخطی بیان می‌شوند. برای جزئیات بیش تر می‌توان به [۵۶] اشاره کرد.

تعریف ۱.۳.۱. *ترانهاده ماتریس A با نماد A^t نشان داده می شود و ماتریسی است که سطرهای آن ستون‌های A و ستون‌های آن سطرهای A هستند، بنابراین اگر $A_{n\times m} = [a_{ij}]$، آن گاه*

$$A^t_{m\times n} = [a_{ji}].$$

هرگاه $n = m$ باشد، ماتریس را مربعی می‌گویند.

تعریف ۲.۳.۱. *ماتریس مربعی که کلیه درایه‌های خارج از قطر اصلی آن صفر باشد یک ماتریس قطری نامیده می‌شود و با نماد $diag(a_{11}, \dots, a_{nn})$ نشان داده می شود.*

تعریف ۳.۳.۱. *ماتریس مربعی را متقارن گویند هرگاه $A = A^t$.*

تعریف ۴.۳.۱. *اگر A یک ماتریس $m \times n$ باشد، بعد فضای سطری یا ستونی A را رتبه ماتریس A نامیده و با $rank(A)$ نشان می دهند.*

تعریف ۵.۳.۱. *یک ماتریس مربعی با رتبه کامل را ناویژه گویند. یک ماتریس ناویژه A دارای معکوس منحصر به فرد است که با A^{-1} نشان داده می‌شود و در شرط $AA^{-1} = A^{-1}A = I$ صدق می‌کند.*

تعریف ۶.۳.۱. *اگر برای ماتریس متقارن $A_{n \times n}$ به ازای تمام مقادیر ممکن $U_{n \times 1}$، غیر از $U = 0$، $U^tAU > 0$ باشد، صورت درجه دوم U^tAU را معین مثبت و A را ماتریس معین مثبت گویند. همچنین اگر $U^tAU \geq 0$، آن‌گاه U^tAU و A را نیمه معین مثبت گویند.*

تعریف ۷.۳.۱. *اگر A ناویژه باشد آن‌گاه A^t نیز ناویژه است. همچنین اگر A و B ناویژه با ابعاد یکسان باشند، آن‌گاه AB نیز ناویژه است.*

اگر A یک ماتریس $p \times q$ با رتبه p و B یک ماتریس $q \times s$ با رتبه q باشد، آنگاه رتبه ماتریس AB برابر است با p. اگر A یک ماتریس $m \times n$ از رتبه r باشد، در این‌صورت A یک تجزیه کامل رتبه‌ای چون $A = FG$ دارد که در آن F و G به ترتیب ماتریس‌هایی $m \times r$ و $r \times n$ هر دو از رتبه r هستند. [۴]

۱.۳.۱ مشتق تابع حقیقی مقدار نسبت به ماتریس

مشتق تابع $f(V)$ نسبت به ماتریس $V_{n\times p} = (V_{ij})$، ماتریسی $n \times p$ است که درایه عمومی آن $\frac{\partial f(V)}{\partial V_{ij}}$ است. یعنی

$$\frac{\partial f(V)}{\partial V} = \left(\frac{\partial f(V)}{\partial V_{ij}}\right).$$

بر اساس تعریف فوق اگر $U = (U_۱, U_۲, \dots, U_p)^t$ آن‌گاه

$$\frac{\partial f(U)}{\partial U^t} = \left(\frac{\partial f(U)}{\partial U_۱}, \dots, \frac{\partial f(U)}{\partial U_p}\right),$$

برداری $p \times ۱$ است ولی $\frac{\partial f(U)}{\partial U}$ بردار $۱ \times p$ است. خواص مهم مشتق تابع حقیقی مقدار نسبت به ماتریس به شرح زیر است:

۱. اگر $V_{p\times p}$ نامتقارن باشد و $|V| \neq ۰$،

$$\frac{\partial}{\partial V}|V| = |V|\left(V^{-۱}\right)^t.$$

۲. اگر $V_{p\times p}$ متقارن باشد،

$$\frac{\partial}{\partial V}|V| = ۲|V|V^{-۱} - diag\left(|V|V^{-۱}\right).$$

۳. برای ماتریس مربع $V_{p\times p}$،

$$\frac{\partial}{\partial V}|V|^{\alpha} = \alpha|V|^{\alpha-۱}\frac{d}{dV}|V|.$$

۴. برای توابع حقیقی مقدار f و g و ماتریس V،

$$\frac{\partial}{\partial V}\left[f(V)g(V)\right] = f(V)\frac{\partial g(V)}{\partial V} + g(V)\frac{\partial f(V)}{\partial V}.$$

۵. برای بردار $U_{p\times ۱}$ و ماتریس $A_{p\times p}$،

$$\frac{\partial}{\partial U}(U^tAU) = ۲AU.$$

۶. برای بردارهای $U_{p\times 1}$ و $a_{p\times 1}$،

$$\frac{\partial}{\partial U}(a^tU) = a.$$

برای ملاحظات بیشتر و اثبات‌ها به [۵۶] می‌توان اشاره نمود.

۲.۳.۱ مشتق بردار نسبت به بردار

مشتق بردار سطری $U^t = (U_1, U_2, \ldots, U_n)$ نسبت به بردار ستونی $V = (V_1, V_2, \ldots, V_p)^t$، ماتریسی $p \times n$ است که درایه سطر iام و ستون jام آن $\left(\frac{\partial U_j}{\partial V_i}\right)_{\substack{i=1,\ldots,p\\ j=1,\ldots,n}}$ است. به عبارت دیگر

$$\frac{\partial U^t}{\partial V} = \left(\frac{\partial U_j}{\partial V_i}\right) = \begin{bmatrix} \frac{\partial U_1}{\partial V_1} & \cdots & \frac{\partial U_n}{\partial V_1} \\ \vdots & & \vdots \\ \frac{\partial U_1}{\partial V_p} & \cdots & \frac{\partial U_n}{\partial V_p} \end{bmatrix} = \left(\frac{\partial U_1}{\partial V} \cdots \frac{\partial U_n}{\partial V}\right).$$

با این تعریف ملاحظه می شود که $\frac{\partial U_j}{\partial V}$ ستون jام ماتریس $\frac{\partial U^t}{\partial V}$ است.

مشتق بردار ستونی U نسبت به بردار ستونی V نیز به طریق مشابه تعریف می‌شود. می‌توان نوشت:

$$\frac{\partial U^t}{\partial V} = \frac{\partial U}{\partial V} = \left(\frac{\partial U_j}{\partial V_i}\right).$$

بنابراین مشتق U نسبت به V، همچون U^t نسبت به V، ماتریسی $p \times n$ با درایه عمومی $\frac{\partial U_j}{\partial V_i}$ است.

۳.۳.۱ ژاکوبی تبدیل‌های ماتریس

اگر تعداد درایه‌های مجزای ماتریس U برابر با درایه‌های مجزای ماتریس V باشند و درایه‌های مجزای U را با $U_1, \ldots, U_p$ و درایه‌های مجزای $\boldsymbol{V}$ با $V_1, \ldots, V_p$ نشان داده شوند، آنگاه ژاکوبی تبدیل $\boldsymbol{V} = g(\boldsymbol{U})$ که آن را ژاکوبی تبدیل $\boldsymbol{U}$ به $\boldsymbol{V}$ می‌نامند و با نماد $j(\boldsymbol{U} \to \boldsymbol{V})$ نشان می‌دهند،