Sari Agriculture Sciences and Natural Resources University

Research on Animal Production, Vol. 16, Issue 4, 2025 p: 1-8

OPEN ACCESS

Research Paper

A Study on the Fitting of a Linear Regression Model for the Live Weight of Biometric Traits using Principal Component Analysis in One-Hump Semnani Camels

Hossein MoradiShahrBabak[1], Hadi Moghbeli[2], Mohmmad MoradiShahrbabak[3], and Milad GholamiTahoone[4]

1- Associate Professor, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran
(Corresponding author: hmoradis@ut.ac.ir)
2- M.Sc. Graduate, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran
3- Professor, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran
4- Ph.D. Student, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran

Received: 10 April, 2025 Revised: 20 July 2025 Accepted: 22 August 2025

Extended Abstract

Background: Due to the continuous and intermittent droughts in the central desert and its outskirts, only camels are capable of survival and production, and other livestock are not able to continue living in most of these areas under these conditions. Given the declining trend of the camel population in Iran, extensive studies are necessary to select suitable animals, and the implementation of breeding programs is particularly necessary. The first step in breeding programs and animal selection is to identify important economic traits and to accurately measure or estimate these traits using correct and accurate methods. Due to the location and spatial conditions of camel life, it is not possible in most cases to access scales for weighing. Since the age of one year is the age of selection among camel breeders for remaining in the herd or sale, the importance of having a correct equation of the dependence of weight on biometric traits at this age becomes more apparent than ever. Therefore, the need to provide an appropriate method in this regard has led to the use of external body measurements to evaluate live camels in terms of weight characteristics. The mechanisms involved in the control of most biological traits in living organisms are highly complex for interpretation by univariate analysis. Because most traits are biologically correlated through the pleiotropic effects of genes and the linkage of gene loci, this correlation causes estimates to be biased; as a result, multivariate methods are used to analyze data with high correlation. The principal component analysis (PCA) method is a mathematical method for converting a number of correlated variables into a smaller number of uncorrelated variables. Using PCA makes it possible to predict the exact size of the desired trait that cannot be measured while measuring a series of traits related to the desired trait. The present study mainly aims to compare the PCA method with the multivariate linear regression and ordinary least squares methods, and ultimately to present a suitable, simple, accurate, and efficient equation for estimating live weight from biometric traits of one-year-old camels.

Methods: In this study, 250 one-year-old camels were randomly selected from different regions of Semnan Province, and their biometric traits and weights were recorded after recording their gender and age. The probability of correlation between two groups was determined using the coefficient of determination (R^2) formula. The variance inflation factor (VIF) formula was used to detect the intensity of multicollinearity in ordinary least squares regression analysis. Data were edited with Excel software. Multiple regression analysis and PCA were performed with SAS software. The presence of multiple alignment in the data was examined by examining the correlation between independent variables. After estimating the relevant eigenvalues and vectors, a number of eigenvalues were used to justify the most changes to create a new regression equation. The Kaiser-Guttman rule was used to select the number of eigenvalues. In the final stage, the coefficients of the regression equation were obtained with standardized data.

Results: The body weight, body height, body length, abdominal circumference, and chest circumference were the traits measured and estimated with mean values of 170 kg, 145, 116, 174, and 134 cm, respectively, and with an $R^2$of 0.92. The presence of a high R^2could be due to variance inflation resulting from collocation and one of the earliest signs of the existence of multiple collocation, which was determined by examining the high correlation between most variables. The variance inflation values for body height, body length, abdominal circumference,

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]

Copyright ©2025 MoradiShahrBabak et al. Published by Sari Agricultural Sciences and Natural Resources University.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 Unported License which allows users to read, copy, distribute and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

and chest circumference as the independent variables were estimated at 45.66, 0.39, 0.92, and 2.05, respectively, with the body height trait having the highest variance compared to the critical value of 5 to 10, where multiple collocation would occur. According to the results of the ordinary least squares analysis, the body height, body length, abdominal circumference, and chest circumference were important traits at the levels of 0.50, 0.49, 0.48, and 0.51 in estimating body weight, respectively, with the chest circumference and abdominal circumference traits having the highest and lowest importance, respectively. According to the PCA results, the body height, body length, abdominal circumference, and chest circumference were important traits at the levels of 0.95, 0.62, 0.76, and 0.78 in estimating body weight, respectively, with the body height and body length traits having the highest and lowest importance, respectively.

Conclusion: The results of this study show that the multiple correlation problem in the data related to the relationship between the weight of Semnani camels and four independent variables related to this trait can be solved using the PCA method. In the case of multiple correlation in the data, the PCA method is recommended for estimating the regression equation compared to the ordinary least squares method. In this case, the standard errors of the estimates will be lower than those the ordinary least squares method. Finally, the obtained results suggest that PCA can be used in the presence of alignment in multivariate linear regression analyses and has more accurate estimates than ordinary least squares methods. Moreover, it can be used with an effective reduction in the number of biometric traits required for use in breeding programs and the selection of superior individuals.

Keywords: Camel, Co-Linearity, Ordinary least squares, Principal components analysis and Trait

How to Cite This Article: MoradiShahrBabak, H., Moghbeli, H., & MoradiShahrbabak, M., GholamiTahoone, M. (2025). A Study on the Fitting of a Linear Regression Model for the Live Weight of Biometric Traits using Principal Component Analysis in One-Hump Semnani Camels. *Res Anim Prod,* 16(4), 1-8. DOI: 10.61882/rap.2025.192

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]

دانشگاه علوم کشاورزی و منابع طبیعی ساری

مقاله پژوهشی

مطالعه برازش مدل‌های تابعیت وزن زنده از صفات بیومتری با تحلیل مؤلفه‌های اصلی در شترهای یک‌کوهانه نژاد سمنانی

حسین مرادی شهربابک[1]، هادی مقبلی[2]، محمد مرادی شهربابک[3] و میلاد غلامی طاحونه[4]

۱- دانشیار، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران، (نویسنده مسوول: hmoradis@ut.ac.ir)
۲- دانش‌آموخته کارشناسی ارشد، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران
۳- استاد، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران
۴- دانشجوی دکتری، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران

تاریخ دریافت: ۱۴۰۴/۰۱/۲۱ تاریخ ویرایش: ۱۴۰۴/۰۴/۲۹ تاریخ پذیرش: ۱۴۰۴/۰۵/۳۱
صفحه ۱ تا ۸

چکیده مبسوط

مقدمه و هدف: با توجه به وجود خشک‌سالی‌های ممتد و متناوب در نواحی کویر مرکزی و حواشی آن، فقط شتر قابلیت زیست و تولید را دارد و سایر دام‌ها قادر به ادامه حیات دراکثر این مناطق با این شرایط نیستند. با توجه به روند کاهشی جمعیت شتر در کشور، انجام مطالعات گسترده جهت انتخاب حیوانات مناسب ضروری است و اجرای برنامه‌های اصلاحی ضرورت خاصی را دارد. اولین مرحله در برنامه‌های اصلاحی و انتخاب حیوانات، شناسایی صفات مهم اقتصادی و اندازه‌گیری و یا برآورد دقیق این صفات با استفاده از روش‌های صحیح و دقیق است. به‌دلیل موقعیت و شرایط مکانی زندگی شتر، در اکثر موارد دسترسی به باسکول جهت وزن‌کشی مقدور نیست و همچنین از آنجا که در بین پرورش دهندگان شتر، سن یک سالگی سن انتخاب برای ماندن در گله یا فروش است، اهمیت داشتن معادله‌ای صحیح از تابعیت وزن از صفات بیومتری در این سن بیش از پیش آشکار می‌شود. بنا بر این، ضرورت ارائه روش مناسب در این خصوص سبب گردیده است تا با استفاده از اندازه‌های ظاهری بدن، جهت ارزیابی شتران زنده از نظر خصوصیات وزن استفاده شود. سازوکارهای درگیر در کنترل اغلب صفات بیولوژیک در موجودات زنده پیچیده‌تر از آن هستند که به‌وسیله آنالیز تک‌متغیره تفسیر شوند. از آنجا که اغلب صفات به‌طور بیولوژیک از طریق اثرات پلیوتروپیک ژن‌ها و پیوستگی جایگاه‌های ژنی با یکدیگر همبستگی دارند و این همبستگی سبب اریب‌شدن برآوردها می‌شود، در نتیجه روش‌های چندمتغیره جهت آنالیز داده‌ها با همبستگی بالا به‌کار گرفته می‌شوند. روش آنالیز مؤلفه‌های اصلی یک روش ریاضی جهت تبدیل تعدادی از متغیرهای همبسته به تعداد کمتری از متغیرهای غیرهمبسته است که در صورت اندازه‌گیری یک سری از صفات مرتبط با صفت مورد نظر، اندازه دقیق صفت مورد نظر که امکان اندازه‌گیری ندارد را می‌توان با استفاده از آنالیز مؤلفه‌های اصلی پیش‌بینی کرد. هدف اصلی پژوهش حاضر، مقایسه روش تحلیل مولفه‌های اصلی (PCA) با روش رگرسیون خطی چندمتغیره و حداقل مربعات معمولی است و در نهایت معادله‌ای مناسب، ساده، دقیق و کارآمد جهت برآورد وزن زنده از صفات بیومتری شترهای یک ساله ارائه می‌گردد.

مواد و روش‌ها: در این پژوهش، به‌صورت تصادفی ۲۵۰ نفر از شترهای یک ساله از مناطق مختلف استان سمنان انتخاب گردیدند که پس از ثبت جنسیت و سن، صفات بیومتری و اوزان ثبت شدند. به‌منظور تعیین میزان احتمال همبستگی میان دو دسته از فرمول ضریب تبیین (R^2) و به‌منظور تشخیص شدت چندخطی در تحلیل رگرسیون حداقل مربعات معمولی از فرمول ضریب تورم واریانس (VIF) استفاده شد. ویرایش داده‌ها با نرم‌افزار Excel و آنالیزهای تابعیت چندگانه و PCA با نرم‌افزار SAS انجام شدند. سپس، وجود هم‌راستایی چندگانه در داده‌ها با بررسی همبستگی بین متغیرهای مستقل بررسی شد. پس از برآورد مقادیر و بردارهای ویژه مربوطه، تعدادی از مقادیر ویژه که بیشترین تغییرات را توجیه می‌کنند برای ایجاد معادله تابعیت جدید مورد استفاده قرار گرفتند. به‌منظور انتخاب تعداد مقادیر ویژه از قانون کایسر–گاتمن استفاده شد. در مرحله آخر، ضرایب معادله تابعیت با داده‌های استاندارد شده به‌دست آمدند.

یافته‌ها: صفات وزن بدن، ارتفاع بدن، طول بدن، محیط شکم و محیط سینه به‌ترتیب با مقادیر میانگین ۱۷۰ کیلوگرم، ۱۴۵، ۱۱۶، ۱۷۴ و ۱۳۴ سانتی‌متر و با ضریب تبیین ۰/۹۲ اندازه‌گیری و برآورد شدند. وجود ضریب تبیین بالا می‌تواند ناشی از تورم واریانس حاصل از هم‌راستایی و یکی از ابتدایی‌ترین نشانه‌های وجود هم‌راستایی چندگانه باشد که با بررسی بین اکثر متغیرها همبستگی بالایی مشخص گردید. مقادیر تورم واریانس برای متغیرهای مستقل ارتفاع بدن، طول بدن، محیط شکم و دور سینه به‌ترتیب ۴۵/۶۶، ۰/۳۹، ۰/۹۲ و ۲/۰۵ برآورد شدند که صفت ارتفاع بدن دارای بیشترین واریانس نسبت به مقدار بحرانی پنج تا ۱۰ بود که هم‌راستایی چندگانه رخ خواهد داد. طبق نتیجه آنالیز روش حداقل مربعات معمولی، صفات ارتفاع بدن، طول بدن، محیط شکم و محیط سینه به‌ترتیب به میزان‌های ۰/۵۰، ۰/۴۹، ۰/۴۸ و ۰/۵۱ در برآورد وزن بدن اهمیت داشتند، که صفات محیط سینه و محیط شکم به‌ترتیب بیشترین و کمترین اهمیت را داشتند. با توجه به نتایج آنالیز تابعیت مولفه‌های اصلی، صفات ارتفاع بدن، طول بدن، محیط شکم و محیط سینه به‌ترتیب به مقادیر ۰/۹۵، ۰/۶۲، ۰/۷۶ و ۰/۷۸ در برآورد وزن بدن مورد اهمیت بودند که صفات ارتفاع بدن و طول بدن به‌ترتیب بیشترین و کمترین اهمیت را داشتند.

نتیجه‌گیری: نتایج حاصل از این مطالعه نشان می‌دهند که مشکل هم‌راستایی چندگانه موجود در اطلاعات مربوط به ارتباط بین وزن شتر نژاد سمنانی با چهار متغیر مستقل مربوط به این صفت با استفاده از روش تابعیت مولفه‌های اصلی قابل حل است. در شرایط وجود هم‌راستایی چندگانه در داده‌ها، روش تابعیت مؤلفه‌های اصلی برای برآورد معادله تابعیت نسبت به روش حداقل مربعات معمولی پیشنهاد می‌شود. در این صورت، خطاهای استاندارد برآوردها نسبت به روش حداقل مربعات معمولی کمتر خواهد بود. در نهایت، نتایج به‌دست‌آمده پیشنهاد می‌کنند که آنالیز مؤلفه‌های اصلی را می‌توان در حالت وجود هم‌راستایی در آنالیزهای رگرسیون خطی چندمتغیره به‌کار برد و برآوردهای دقیق‌تری نسبت به روش‌های حداقل مربعات معمولی داشت. همچنین، این روش را با کاهش مؤثر در تعداد صفات بیومتریک مورد نیاز می‌توان برای استفاده در برنامه‌های اصلاحی و انتخاب افراد برتر استفاده کرد.

واژه‌های کلیدی: تحلیل مؤلفه‌های اصلی، حداقل مربعات معمولی، شتر، صفت و هم‌راستایی چندگانه

مقدمه

کشور ایران دارای ۸۳/۳ میلیون هکتار زمین مرتعی است که از نظر پوشش‌گیاهی ۶/۵۱ درصد آن دارای وضعیت خوب، ۲۵/۱۲ درصد آن دارای وضعیت متوسط و ۶۸/۳۷ درصد آن دارای وضعیت ضعیف هستند (Agricultural statistics, 2022). با توجه به وجود خشک‌سالی‌های ممتد و متناوب در نواحی کویر مرکزی و حواشی آن، فقط شتر قابلیت زیست و تولید را دارد و سایر دام‌ها قادر به ادامه حیات در اکثر این مناطق با این شرایط نیستند (Shokri, 1997). شتر با تولیدات و خدمات فراوان شامل پشم، گوشت، شیر، حمل و نقل و غیره

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]

در خدمت بشر است. در ایران، سه گروه شتر دوکوهانه، یک‌کوهانه و آمیخته وجود دارند (Nazeradl, 1986). با توجه به روند کاهشی جمعیت شتر در کشور، انجام مطالعات گسترده جهت انتخاب حیوانات مناسب ضروری است و اجرای برنامه‌های اصلاحی ضرورت خاصی را دارد. اولین مرحله در برنامه‌های اصلاحی و انتخاب حیوانات، شناسایی صفات مهم اقتصادی و اندازه‌گیری و یا برآورد دقیق این صفات با استفاده از روش‌های صحیح و دقیق است. به‌دلیل موقعیت و شرایط مکانی زندگی شتر، دسترسی به باسکول جهت وزن‌کشی در اکثر موارد مقدور نیست و همچنین از آنجا که در بین پرورش‌دهندگان شتر، سن یک سالگی سن انتخاب برای ماندن در گله یا فروش است، اهمیت داشتن معادله‌ای صحیح از تابعیت وزن از صفات بیومتری در این سن بیش از پیش آشکار می‌شود. بنا بر این، ضرورت ارائه روش مناسب در این خصوص سبب گردیده است تا از اندازه‌های ظاهری بدن جهت ارزیابی شتران زنده از نظر خصوصیات وزن استفاده شوند. مکانیسم‌های درگیر در کنترل اغلب صفات بیولوژیک در موجودات زنده پیچیده‌تر از آن هستند که به‌وسیله آنالیز تک‌متغیره تفسیر شوند زیرا اغلب صفات به‌طور بیولوژیک از طریق اثرات پلیوتروپیک ژن‌ها و پیوستگی جایگاه‌های ژنی با یکدیگر همبستگی دارند و این همبستگی سبب اریب‌شدن برآوردها می‌شود (Price *et al*., 2006). روش‌های چندمتغیره جهت آنالیز داده‌ها با همبستگی بالا به‌کار گرفته می‌شوند (Somers *et al*., 1986). اگر بین متغیرهای مستقل در تابعیت چندمتغیره رابطه خطی وجود داشته‌باشد، گفته می‌شود که نامتعامد هستند. وقتی متغیرهای مستقل متعامد باشند، به آسانی می‌توان از تابعیت چندمتغیره استفاده کرد. متأسفانه، متغیرهای مستقل در اکثر موارد متعامد نیستند. گاهی متعامد نبودن، مشکلاتی را فراهم نمی‌سازد اما در برخی موارد متغیرهای مستقل همبستگی دارند و در چنین مواردی استنتاج بر مبنای مدل تابعیت می‌تواند گمراه‌کننده باشد. در حالتی که بین متغیرهای مستقل وابستگی خطی مشاهده شود، گفته می‌شود که هم‌راستایی چندگانه وجود دارد (Myers *et al*., 1990).

روش آنالیز مؤلفه‌های اصلی یک روش ریاضی جهت تبدیل تعدادی از متغیرهای همبسته به تعداد کمتری از متغیرهای غیرهمبسته است که مؤلفه‌های اصلی نامیده می‌شوند (Johnson *et al*., 2002). در واقع، این روش آنالیز باعث کاهش متغیرها می‌شود و هم‌خطی بین متغیرهای مستقل، که منجر به تفسیر اشتباه از این قبیل داده‌ها می‌شود، را از بین می‌برد (Baker *et al*., 1988). در واقع، در صورت اندازه‌گیری یک سری از صفات مرتبط با صفت مورد نظر، اندازه دقیق صفت مورد نظر که امکان اندازه‌گیری ندارد را می‌توان با استفاده از آنالیز مؤلفه‌های اصلی پیش‌بینی کرد. فریتز و همکاران (1971) برای اولین‌بار روش تابعیت مؤلفه‌های اصلی را برای برآورد ضرایب تابعیت در داده‌هایی با مشکل هم‌راستایی چندگانه ارائه کردند. پژوهشگران زیادی از روش آنالیز مؤلفه‌های اصلی جهت تخمین وزن بدن (Baker *et al*., 1988) و صفات عملکردی (Blackith *et al*., 1971) استفاده کردند. آنالیز مؤلفه‌های اصلی جهت توصیف همبستگی بین اندازه‌گیری‌های ظاهری مربوط به بدن و اندازه بدن در جوجه‌های گوشتی (Hocking *et al*., 1996; Lehmann *et al*., 2006) و بوقلمون (Sakomura *et al*., 2011) استفاده شده‌است. این تکنیک همچنین برای کاهش تعداد متغیرهای مستقل جهت پیش‌بینی ارزش‌های اصلاحی ژنومیکی استفاده شد (Price *et al*., 2006).

هدف اصلی پژوهش حاضر، مقایسه روش تحلیل مؤلفه‌های اصلی[1] (PCA) با روش رگرسیون خطی چندمتغیره و حداقل مربعات معمولی بود و در نهایت معادله‌ای مناسب، ساده، دقیق و کارآمد جهت برآورد وزن زنده از صفات بیومتری شترهای یک‌ساله ارائه گردید.

مواد و روش‌ها

در این پژوهش، به‌صورت تصادفی ۲۵۰ نفر از شترهای یک‌ساله از مناطق مختلف استان سمنان انتخاب گردیدند. پس از ثبت جنسیت و سن صفات بیومتری (ارتفاع از جدوگاه، محیط سینه، محیط شکم و طول) با متر پارچه‌ای و کولیس با دقت یک سانتی‌متر و اوزان با استفاده از باسکول با دقت ۱۰۰ گرم ثبت شدند.

به‌منظور تعیین میزان احتمال همبستگی میان دو دسته از فرمول ضریب تبیین (R^2) و به‌منظور تشخیص شدت چندخطی در تحلیل رگرسیون حداقل مربعات معمولی از فرمول ضریب تورم واریانس (VIF) طبق فرمول‌های زیر استفاده شد:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y-\hat{Y})^2}{\sum_{i=1}^{n}(Y-\bar{Y})^2}$$

$$\mathrm{VIF} = \frac{1}{(1-R^2)}$$

ویرایش داده‌ها با نرم‌افزار Excel و آنالیزهای تابعیت چندگانه و PCA با رویه GLM و PRINCOM نرم‌افزار SAS انجام شدند (Smith *et al*., 2002). وجود هم‌راستایی چندگانه در داده‌ها، با بررسی همبستگی بین متغیرهای مستقل بررسی شد. به‌منظور انجام روش تابعیت مؤلفه‌های اصلی، ابتدا داده‌ها طبق فرمول زیر استاندارد شدند:

$$x_{ij}^{*} = \frac{x_{ij} - \bar{x}_i}{s_i}$$

که در آن، x_{ij}^{*} متغیر مستقل استانداردشده و $\bar{x}_i$ و s_i به‌ترتیب میانگین و انحراف‌معیار متغیر مستقل مربوطه هستند. پس از برآورد مقادیر و بردارهای ویژه مربوطه، تعدادی از مقادیر ویژه که بیشترین تغییرات را توجیه می‌کنند برای ایجاد معادله تابعیت جدید مورد استفاده قرار گرفتند. به‌منظور انتخاب تعداد مقادیر ویژه از قانون کایسر–گاتمن (Kaiser, 1992) استفاده شد. در این روش، مقادیر ویژه بزرگ‌تر از یک برای ایجاد معادله تابعیت استفاده می‌شوند. به‌منظور انتخاب تعداد مقادیر ویژه مورد نیاز برای انجام تابعیت مولفه‌های اصلی از

[1] Principal components analysis

[Downloaded from rap.sanru.ac.ir on 2026-11-13]
[DOI: 10.66224/rap.2025.192]

فرمول $\frac{\sum_{j=1}^{k} \lambda_j}{k} > 0.99$ استفاده شد و در نتیجه مولفه ۱ از ۴ مولفه موجود مورد استفاده قرارگرفت. معادله جدید برآورد ضرایب هر کدام از بردارهای ویژه مورد استفاده در زیر نشان داده شده است (Bryant *et al*., 1995; Jolliffe, 2005; Raskin *et al*., 1988):

$$y = b_0 + b_1(prin\ 1) + \ldots + b_k(prin\ k) + e$$

که در آن، b برابر ضرایب تابعیت b_0 برابر عرض از مبداء، prinها بردارهای ویژه، k تعداد مقادیر ویژه تعیین‌شده و y متغیر وابسته استاندارد شده هستند. در مرحله آخر، ضرایب معادله تابعیت با داده‌های استاندارد شده به‌دست آمدند (Raskin & Terry, 1988).

نتایج و بحث

خلاصه آماری صفات مورد بررسی و نتایج مربوط به ضرایب تابعیت مربوط به هرکدام از صفات و نیز خطای استاندارد مربوط به هر ضریب با روش حداقل مربعات معمولی به‌دست آمدند (جداول ۱ و ۲).

جدول ۱- خلاصه آماری صفات اندازه‌گیری‌شده

Table 1. A statistical summary of the measured traits

صفت Trait	کمینه Minimum	بیشینه Maximum	انحراف معیار Standard deviation	میانگین Average
وزن بدن (کیلوگرم) Body weight (kg)	105	218	29.8	170
ارتفاع بدن (سانتی‌متر) Body height (cm)	127	162	7.7	145
طول بدن (سانتی‌متر) Body length (cm)	93	127	8.2	116
محیط شکم (سانتی‌متر) Abdominal circumference (cm)	145	200	14.8	174
محیط سینه (سانتی‌متر) Chest circumference (cm)	103	152	11.1	134

جدول ۲- همبستگی‌های بین متغیرهای مستقل

Table 2. Correlations between independent variables

صفت Trait	وزن بدن (کیلوگرم) Body weight (kg)	ارتفاع بدن (سانتی‌متر) Body height (cm)	طول بدن (سانتی‌متر) Body length (cm)	محیط شکم (سانتی‌متر) Abdominal circumference (cm)	محیط سینه (سانتی‌متر) Chest circumference (cm)
وزن بدن (کیلوگرم) Body weight (kg)	1	0.77	0.66	0.76	0.85
ارتفاع بدن (سانتی‌متر) Body height (cm)	0.77	1	0.62	0.71	0.80
طول بدن (سانتی‌متر) Body length (cm)	0.66	0.62	1	0.75	0.86
محیط شکم (سانتی‌متر) Abdominal circumference (cm)	0.76	0.71	0.75	1	0.88
محیط سینه (سانتی‌متر) Chest circumference (cm)	0.85	0.80	0.86	0.88	1

ضریب تبیین این برآوردها ۰/۹۲ بود؛ وجود ضریب تبیین بالا در این حالت می‌تواند یکی از ابتدایی‌ترین نشانه‌های وجود هم‌راستایی چندگانه باشد. برای این منظور، قدم بعدی کنترل همبستگی‌ها بین متغیرهای مستقل است (Kistler & Wightman, 1992; Waring *et al*., 2010). طبق جدول (۲)، بین اکثر متغیرها همبستگی بالایی وجود دارد؛ در نتیجه، با وجود این همبستگی‌ها باید به‌وجود هم‌راستایی چندگانه مشکوک شد. برآورد ضریب تبیین بالای به‌دست‌آمده در این حالت می‌تواند ناشی از تورم واریانس حاصل از هم‌راستایی باشد، زیرا در این حالت شرط نبود رابطه بین متغیرها برای معکوس ماتریس X'X به‌منظور برآورد پارامترها در روش حداقل مربعات معمولی به‌هم می‌خورد و تورم واریانس ایجاد می‌شود؛ پس به روش بهتری برای برآوردهای دقیق‌تر نیاز است. مرحله بعد، محاسبه معادله تابعیت مورد نظر از طریق روش تابعیت مؤلفه‌های اصلی است که به این منظور ابتدا باید مقادیر ویژه و بردارهای ویژه مربوطه را برآورد نمود. با توجه به جدول (۳)، صفت ارتفاع دارای واریانس بیشتری نسبت به مقدار بحرانی است (مقدار بحرانی ۵ تا ۱۰ است) که در این حالت هم‌راستایی چندگانه رخ خواهد داد. در صورتی که هدف، حذف این متغیرها از مدل نباشد و حضور آن‌ها در مدل ضروری باشد، باید از روش دیگری مانند تابعیت مولفه‌های اصلی استفاده شود که برآوردهایی دقیق‌تر و با خطای استاندارد کمتر است.

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]

جدول ۳- مقادیر تورم واریانس برای متغیرهای مستقل

Table 3. Variance inflation values for independent variables

صفت Trait	تورم واریانس Variance inflation
ارتفاع بدن (سانتی‌متر) Body height (cm)	45.66
طول بدن (سانتی‌متر) Body length (cm)	0.39
محیط شکم (سانتی‌متر) Abdominal circumference (cm)	0.92
محیط سینه (سانتی‌متر) Chest circumference (cm)	2.05

وجود تعداد کمی از مقادیر ویژه با مقدار زیاد نشان می‌دهد که تعداد کمی از متغیرها بیشتر تنوع موجود در متغیرهای مشاهده‌شده را توضیح می‌دهند. مقدار ویژه صفر یعنی هم‌راستایی کامل بین متغیرهای مستقل وجود دارد و مقادیر ویژه بسیار کوچک نشان‌دهنده هم‌راستایی چندگانه قوی هستند. جهت محاسبه معادله تابعیت مورد نظر از روش تابعیت مولفه‌های اصلی ابتدا مقادیر ویژه و بردارهای ویژه مربوطه برآورد شد (جداول ۴ و ۵). طبق جدول (۴)، مقادیر ویژه از بالا به پایین به‌ترتیب کاهش می‌یابند و همچنین با افزایش شماره مقادیر ویژه میزان سهم آن‌ها در توضیح واریانس نیز کاهش می‌یابد. به این ترتیب، مقدار ویژه اول بالاترین سهم از کل واریانس (۰/۷۹) و مقدار ویژه آخر نیز کمترین مقدار از کل واریانس (۰/۰۵) را توضیح می‌دهند. با توجه به این که داده‌ها در این روش استاندارد شدند و دارای میانگین صفر و واریانس یک هستند، پس واریانس کل برابر با تعداد متغیرهایی است که در آنالیز شرکت می‌کنند که در این مطالعه برابر با ۴ است. در نتیجه، به‌عنوان مثال، مقدار ویژه اول ۳/۱۴ است که به‌عبارت دیگر ۳/۱۴ واحد از ۴ واحد واریانس را توضیح می‌دهد و به همین ترتیب برای دیگر مقادیر ویژه نیز برآورد شد. با توجه به جدول (۵)، مقدار ویژه اول بیشترین واریانس را توضیح می‌دهد و بقیه مقادیر ویژه نقش کمتری را دارند که این دلیلی دیگر بر وجود هم‌راستایی چندگانه است.

جدول۴- مقادیر ویژه برای متغیرهای مورد استفاده

Table 4. Eigenvalues for variables used

ردیف Row	واریانس* Variance	مقادیر ویژه Eigenvalues
1	0.79	3.14
2	0.10	0.40
3	0.06	0.23
4	0.05	0.21

* میزان واریانسی که این مقدار ویژه از کل واریانس موجود را توضیح می‌دهد.

جدول ۵- بردارهای ویژه برای متغیرهای مورد استفاده

Table 5. Eigenvectors for the variables used

صفت Trait	بردار وبژه ۱ Prin1	بردار وبژه ۲ Prin2	بردار وبژه ۳ Prin3	بردار وبژه ۴ Prin4
ارتفاع بدن (سانتی‌متر) Body height (cm)	0.50	-0.35	-0.47	-0.62
طول بدن (سانتی‌متر) Body length (cm)	0.49	-0.54	0.63	0.23
محیط شکم (سانتی‌متر) Abdominal circumference (cm)	0.48	0.73	0.37	-0.30
محیط سینه (سانتی‌متر) Chest circumference (cm)	0.51	0.19	-0.48	0.68

با استفاده از تابعیت مؤلفه‌های اصلی، داده‌ها متعامد یا غیرهمبسته شدند و سپس معادله تابعیت مورد نظر به‌دست آمد. به‌منظور انتخاب تعداد مقادیر ویژه مورد نیاز برای انجام تابعیت مؤلفه‌های اصلی طبق فرمول، در نتیجه مؤلفه ۱ از ۴ مؤلفه موجود مورد استفاده قرارگرفت. این برآوردها به‌همراه خطاهای استاندارد در جدول (۶) نشان داده شده‌اند.

جدول ۶- برآوردهای به‌دست‌آمده برای معادله تابعیت با استفاده از تابعیت مولفه‌های اصلی

Table 6. Estimates obtained for the dependency equation using principal component dependency

ردیف Row	برآورد Estimate	خطای استاندارد Standard error
عرض از مبدا y-intercept	170	1.19
بردار ویژه ۱ Prin1	16	0.67
بردار ویژه ۲ Prin2	4.48	1.87
بردار ویژه ۳ Prin3	-0.36	2.45
بردار ویژه ۴ Prin4	-0.96	2.63

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]

در مرحله بعد، باید این برآوردها را به مقادیر اولیه تبدیل کرد. به این منظور، ابتدا این مقادیر باید به ضرایب تابعیت متغیرهای مستقل استانداردشده (برآوردکننده‌های مولفه‌های اصلی) و سپس به ضرایب تابعیت برای متغیرهای مستقل قبل از استانداردشدن تبدیل گرددند. این کار با توجه به معادلات بیان‌شده در قسمت مواد و روش‌ها انجام شد. این برآوردها به‌همراه خطاهای استاندارد مربوطه در جدول (۷) نشان داده شده‌اند. با توجه به نتایج آنالیز تابعیت مولفه‌های اصلی، صفت ارتفاع بیشترین و طول بدن کمترین اهمیت را در تعیین معادله تابعیت برای پیش‌بینی صفت وزن نشان دادند؛ در حالی‌که در روش حداقل مربعات معمولی، صفت دور سینه کمترین اهمیت را در برآورد وزن داشت.

جدول ۷- برآوردهای به‌دست‌آمده برای معادله تابعیت بعد از تبدیل برای متغیرهای ابتدایی

Table 7. Estimates obtained for the dependency equation after transformation for the initial variables

صفت Trait	برآورد Estimate	خطای استاندارد Standard error
عرض از مبدا y-intercept	-278.86	3.2
ارتفاع بدن (سانتی‌متر) Body height (cm)	0.95	0.28
طول بدن (سانتی‌متر) Body length (cm)	0.62	0.22
محیط شکم (سانتی‌متر) Abdominal circumference (cm)	0.76	0.20
محیط سینه (سانتی‌متر) Chest circumference (cm)	0.78	0.12

در پژوهشی دیگر، معادله تابعیت با روش‌های تابعیت رایج و مؤلفه‌های اصلی برای وزن بدن در گاوهای سیاه ژاپنی با استفاده از ۱۳ متغیر مستقل بررسی شد. و نتیجه گرفتند که استفاده از آنالیز مؤلفه‌های اصلی برای برآورد ضرایب تابعیت پایدار و به‌منظور غلبه بر مشکل هم‌راستایی چندگانه مناسب بود (Malau-Aduli *et al*., 2004). همان‌طورکه بیان شد، یکی از خصوصیات روش تابعیت مولفه‌های اصلی برآوردهای با خطای استاندارد کمتر نسبت به روش حداقل مربعات است و این موضوع در این مطالعه نیز تایید شد. خطاهای استاندارد روش تابعیت مولفه‌های اصلی نسبت به روش حداقل مربعات کمتر هستند و اختلاف معنی‌دار دارند. موضوع دیگر در این مورد این است که با توجه به این که در روش حداقل مربعات معمولی متغیرها نامتعامد هستند، در نتیجه برآوردها دقت لازم را ندارند و اریب هستند و ممکن است با اضافه شدن داده‌های دیگر به مجموعه داده‌های فعلی، این برآوردها تغییر کنند. به این ترتیب، این برآوردها برای ایجاد یک معادله تابعیت مناسب نیستند و برآوردهای انجام‌شده با تابعیت مولفه‌ها اصلی که بر مبنای ایجاد داده‌ها متعامد و سپس برآورد پارامترها عمل می‌کند، مناسب‌ترند.

نتیجه‌گیری کلی

نتایج حاصل از این مطالعه نشان می‌دهند که مشکل هم‌راستایی چندگانه موجود در اطلاعات مربوط به ارتباط بین وزن شتر نژاد سمنانی با چهار متغیر مستقل مربوط به این صفت با استفاده از روش تابعیت مولفه‌های اصلی قابل حل است. در شرایط وجود هم‌راستایی چندگانه در داده‌ها، روش تابعیت مولفه‌های اصلی برای برآورد معادله تابعیت نسبت به روش حداقل مربعات معمولی پیشنهاد می‌شود. در این صورت، خطاهای استاندارد برآوردها نسبت به روش حداقل مربعات معمولی کمتر خواهد بود. در نهایت، نتایج به‌دست‌آمده پیشنهاد می‌کنند که آنالیز مؤلفه های اصلی را می‌توان در حالت وجود هم‌راستایی در آنالیزهای رگرسیون خطی چندمتغیره به‌کار برد و برآوردهای دقیق‌تری نسبت به روش‌های حداقل مربعات معمولی داشت. ضمناً این روش را می‌توان با کاهش مؤثر در تعداد صفات بیومتریک مورد نیاز برای استفاده در برنامه‌های اصلاحی و انتخاب افراد برتر استفاده کرد.

References

Agricultural statistics, (2022). Volume Two. 191-192.

Baker, J. F., Stewart, T. S., Long, C. R., & Cartwright, T. C. (1988). Multiple regression and principal components analysis of puberty and growth in cattle. *Journal of Animal Science*, *66*(9), 2147-2158.

Blackith, R. E., & Reyment, R. A. (1971). *Multivariate morphometrics* (pp. ix+-412).

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), Reading and understanding multivariate statistics (pp. 99–136). American Psychological Association.

Hocking, P. M., Maxwell, M. H., & Mitchell, M. A. (1996). Relationships between the degree of food restriction and welfare indices in broiler breeder females. *British Poultry Science*, *37*(2), 263-278.

Johnson, R. A., & Wichern, D. W. (2002). Applied multivariate statistical analysis.

Jolliffe, I. (2005). Principal component analysis: Wiley online library. *Google Scholar*.

Kistler, D. J., & Wightman, F. L. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *The Journal of the Acoustical Society of America*, *91*(3), 1637-1647.

Malau-Aduli, A. E. O., Aziz, M. A., Kojima, T., Niibayashi, T., Oshima, K., & Komatsu, M. (2004). Fixing collinearity instability using principal component and ridge regression analyses in the relationship between body measurements and body weight in Japanese Black cattle. *Journal of Animal and Veterinary Advances*, 3(12), 856-863.

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]

Myers, R. H. (1990). Classical and modern regression with applications (Vol. 2, p. 488). Belmont, CA: Duxbury press.
Nazeradl, k. (1986). Camel breeding. First edition. Publications of the Extracurricular Unit of the Cultural Department of the Central Office of the Academic Jihad.
Pinto, L. F. B., Packer, I. U., De Melo, C. M. R., Ledur, M. C., & Coutinho, L. L. (2006). Principal components analysis applied to performance and carcass traits in the chicken. *Animal Research*, *55*(5), 419-425.
Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904-909.
Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology*, *54*(5), 890.
Raychaudhuri, S., Stuart, J. M., & Altman, R. B. (1999). Principal components analysis to summarize microarray experiments: application to sporulation time series. In *Biocomputing 2000* (pp. 455-466).
Shokri, M. M. (1997). Camel and its breeding. First edition. Nourbakhsh Publications.
Sakomura, N. K., Gous, R. M., Marcato, S. M., & Fernandes, J. B. K. (2011). A description of the growth of the major body components of 2 broiler chicken strains. *Poultry Science*, *90*(12), 2888-2896.
Smith, L. I. (2002). A tutorial on principal components analysis.
Somers, K. M. (1986). Multivariate allometry and removal of size with principal components analysis. *Systematic Biology*, *35*(3), 359-368.
Waring, M. E., Eaton, C. B., Lasater, T. M., & Lapane, K. L. (2010). Correlates of weight patterns during middle age characterized by functional principal components analysis. *Annals of Epidemiology*, *20*(3), 201-209.

[Downloaded from rap.sanru.ac.ir on 2026-11-13]

[DOI: 10.66224/rap.2025.192]