

УДК 519.711.3

М.В. Потанина, ст. преподаватель*Севастопольский национальный технический университет**ул. Университетская 33, г. Севастополь, Украина, 99053**E-mail: serg_potanin@mail.ru***ПОВЫШЕНИЕ РЕНТАБЕЛЬНОСТИ ПРОЦЕССА ГОРОДСКИХ АВТОБУСНЫХ ПАССАЖИРОПЕРЕВОЗОК**

Рассмотрена методика оптимизации процесса городских автобусных пассажироперевозок с точки зрения рентабельности автотранспортного средства путем одновременного использования смещенного оценивания параметров регрессионных моделей и процедуры селекции моделей оценки рентабельности.

Ключевые слова: оптимизация, рентабельность, модель, селекция, регуляризация.

Постановка проблемы. Пассажирский автотранспорт превратился в один из основных и наиболее распространенных видов пассажирского транспорта страны.

Главной проблемой пассажирского транспорта является повышение эффективности его работы. Одним из главных финансовых показателей экономической эффективности предприятия, осуществляющего пассажироперевозки, является рентабельность использования автотранспортных средств.

Для повышения этого показателя необходимо достигнуть нового уровня пассажироперевозок, который позволит сделать сбалансированным количественный и качественный состав участников автомобильного пассажирского движения. Это поможет создать правильную пропорцию между автобусами и маршрутными такси, причем с приоритетом в сторону автобусов, таким образом, оптимизируя весь процесс пассажироперевозок.

Анализ последних исследований и публикаций. Обычно при оценке рентабельности процесса пассажироперевозок учитываются такие его параметры как общая прибыль и общие затраты на содержание автобусов. Но на практике каждое автотранспортное средство характеризуется огромным числом показателей, которые обычно коррелируют между собой, т.е. имеет место сильная мультиколлинеарность экспериментальных данных. Структура модели данного сложного объекта, которым является автотранспортное средство, не определена.

Устранение отрицательных последствий мультиколлинеарности частично решает переход к регуляризованным оценкам. В качестве регуляризации предлагается использование смещенных методов оценивания регрессионных моделей, например, таких как ридж-регрессия. Полученные оценки регрессионных коэффициентов являются более устойчивыми, чем оценки МНК и имеют меньшее значение среднеквадратической ошибки прогноза [1].

Процедура регуляризации была обоснована в целом ряде работ А.Н. Тихонова и В.Я. Арсенина. Независимо от этих работ появились результаты Херла и Кеннарда [2], а также еще целый ряд исследований в данной области.

Но при смещенном оценивании параметров центральной проблемой является выбор параметра регуляризации. Главная трудность заключается в том, что и смещение и квадратическая ошибка прогноза зависят от неизвестных истинных значений параметров.

Нерешенные ранее части общей проблемы. Из подробного анализа публикаций по данной тематике видно, что большинство исследователей подчеркивают важность устранения проблемы мультиколлинеарности и селекции полученных прогнозирующих моделей. Но, несмотря на достаточную освещенность отдельных вопросов, связанных со смещенным оцениваем регрессионных коэффициентов, до сих пор имеется множество нерешенных вопросов, связанных с выбором прогнозирующей модели в подобной ситуации.

Не существует общей рекомендуемой методики построения и выбора модели оценки рентабельности процесса пассажироперевозок, с учетом неопределенности структуры модели и особенностей специфики самого метода смещенного оценивания.

Формулировка цели статьи. В настоящей работе решается задача повышения рентабельности процесса городских автобусных пассажироперевозок за счет его оптимизации. Рассмотрена разработанная автором методика одновременного использования смещенного оценивания параметров регрессионных уравнений и процедуры селекции моделей оценки рентабельности автотранспортного средства. При этом предлагается использовать выбор параметра регуляризации r на основе минимума среднеквадратической ошибки прогноза, полученной по этой модели.

Изложение основного материала. В виде приложения разработанной методики рассматривалась задача получения наилучшей прогнозирующей модели оценки рентабельности при использовании автотранспортного средства среднего класса на разных маршрутах предприятия-перевозчика. В качестве Y выступает показатель рентабельности выбранного транспортного средства. В качестве X использовались

данные, характеризующие это транспортное средство: протяженность маршрута, количество автобусов на маршруте по расписанию, количество автобусов на маршруте по списку, продолжительность обратного рейса на маршруте, эксплуатационная скорость, интервал движения в часы пик и т.д. Все X в рассматриваемом примере, в силу своей специфики, являются сильно коррелированными, количество их достаточно велико, что делает невозможным точное определение тех параметров, которые должны присутствовать в прогнозируемой модели, а также значительно снижает ее прогнозирующие способности.

Данная методика использует преимущество ридж-регрессии работать со смещенными оценками коэффициентов прогнозирующей модели. При этом полученная модель объекта дает наименьшую ошибку прогноза в прогнозируемой области.

В качестве критерия оптимальности регрессионной модели берем величину ошибки прогноза, полученной по этой модели. Будем предполагать, что модель используется для целей прогноза величины y в заданной области значений величины x .

Для любой модели можно определить величину среднего риска [1]:

$$I(\alpha) = \int (y - \varphi(x, \alpha))^2 P(x, y) dx dy, \quad (1)$$

где $P(x, y)$ – совместная плотность распределения переменных x и y . Задача имеет различные решения в зависимости от области прогнозирования.

Пусть цель моделирования заключается в получении оптимальной модели для лучшего предсказания функции отклика y в области W_1 , тогда функцию потерь для модели j можно будет записать в виде [3,4]:

$$L(j) = \int_{W_1} E(y - \eta_j(x, \alpha_j))^2 dx / \int_{W_1} dx, \quad (j=1, 2, \dots, q) \quad (2)$$

Здесь вычисляется среднеквадратическая ошибка прогноза для каждой модели j области W_1 , y определяется как возможная величина зависимой переменной в точке x . Если дисперсия σ^2 является константой, то функцию потерь для модели можно записать в виде:

$$L(j) = \sigma^2 + \int_{W_1} (\eta(x) - \eta_j(x, \alpha_j))^2 dx / \int_{W_1} dx \quad (3)$$

Для того, чтобы выбрать аппроксимирующую модель, необходимо иметь значения экспериментальных данных в каждой точке $x_i^T = (x_{i1}, x_{i2}, \dots, x_{ki})$ ($i=1, 2, \dots, N$). Пусть матрица $D^T = [x_1, x_2, \dots, x_N]$ – матрица экспериментальных данных размера $N \times k$ и $Y^T = [y_1, y_2, \dots, y_N]$ – вектор N наблюдений над функцией отклика.

Пусть

$$\hat{\alpha}_j = \Phi_j(X, Y) \quad (4)$$

будет оценкой α_j , где Φ_j функция от Y , плана D и класса моделей S_j .

Если модель линейная по параметрам и используется метод ридж-регрессии [2], тогда

$$\hat{\alpha}_j = (F_j^T F_j + r I_n)^{-1} F_j^T Y, \quad (5)$$

где $F_j^T = [f_j(x_1), f_j(x_2), \dots, f_j(x_N)]$, матрица значений вектора f_j в N экспериментальных точках, r – малое положительное число, добавленное к диагональным элементам матрицы $F_j^T F_j$, так называемый коэффициент регуляризации, а I_n – единичная матрица. Предполагается, что матрица F_j имеет полный ранг.

Рассмотрим выражение для функции потерь в случае ридж-регрессии для полной модели q по области W_1 , где $x \in [-1, 1]$.

$$\begin{aligned} L(q) &= E \int_{W_1} x^T (\alpha - \hat{\alpha}(r)) (\alpha - \hat{\alpha}(r))^T x dx / \int_{W_q} dx = \\ &= \sigma^2 sp X B J B X^T + \alpha^T (I_q - X^T X B) J (I_q - B X^T X) \alpha, \end{aligned} \quad (6)$$

где E – математическое ожидание, $B = (X^T X + r I_q)^{-1}$ и $J = \int_{W_1} x x^T dx / \int_{W_1} dx$.

Формула (6) состоит из двух частей. Первая, содержащая истинные коэффициенты $\alpha^T (I_q - X^T X B) J (I_q - B X^T X) \alpha$, это есть влияние смещения оценок, а вторая часть $\sigma^2 sp X B J B X^T$ содержит дисперсию ошибки [3, 4].

Очевидно, что можно найти такое значение параметра регуляризации, при котором можно существенно уменьшить квадратическую ошибку оценок за счет их небольшого смещения. Автором статьи разработана итерационная процедура поиска оптимального параметра регуляризации r в условиях мультиколлинеарности и определения неизвестных параметров регрессионной модели [3, 4].

В результате этого удалось оптимизировать первоначальную структуру модели, правильно произвести ее спецификацию.

Модель рентабельности пассажироперевозок в общем виде будет иметь вид:

$$Y = Ef\{Y_{\text{Э}}, Y_{\text{Б}}, Y_{\text{М}}\}, \quad (7)$$

где E – математическое ожидание, $Y_{\text{Э}}$ – рентабельность автобусов марки «Эталон», $Y_{\text{Б}}$ – рентабельность автобусов марки «Богдан», $Y_{\text{М}}$ – рентабельность автобусов марки «Мерседес спринтер».

По реальным данным, для трех марок автобусов «Эталон», «Богдан» и «Мерседес спринтер» были составлены матрицы $X_{\text{Э}}, X_{\text{Б}}, X_{\text{М}}$ размерностью 100×30 .

Тогда, в соответствии с марками, рентабельность каждого автотранспортного средства в общем виде будет иметь вид:

$$\begin{aligned} Y_{\text{Э}} &= \alpha_0 + \alpha_1 x_{1\text{Э}} + \alpha_2 x_{2\text{Э}} + \alpha_3 x_{3\text{Э}} + \dots + \alpha_{30} x_{30\text{Э}} + \varepsilon_1 \\ Y_{\text{Б}} &= \alpha_0 + \alpha_1 x_{1\text{Б}} + \alpha_2 x_{2\text{Б}} + \alpha_3 x_{3\text{Б}} + \dots + \alpha_{30} x_{30\text{Б}} + \varepsilon_2 \\ Y_{\text{М}} &= \alpha_0 + \alpha_1 x_{1\text{М}} + \alpha_2 x_{2\text{М}} + \alpha_3 x_{3\text{М}} + \dots + \alpha_{30} x_{30\text{М}} + \varepsilon_3 \end{aligned} \quad (8)$$

Здесь в качестве $x_{i\text{Э}}, x_{i\text{Б}}, x_{i\text{М}}$ выступают параметры, характеризующие каждую марку автобусов. С помощью разработанной автором методики были построены прогнозирующие модели оценки рентабельности. Для каждой марки автобуса была рассчитана рентабельность в определенный период времени по полной модели. Все x_i были стандартизированы, Y – нет, дисперсия единичная. Было разработано программное обеспечение на языке m-файлов с помощью программы MATLAB6p1. Очевидно, что рентабельность в разное время суток для разных марок автобусов очень варьирует. Традиционно высокая в часы пик, очень низкая в вечернее время. Для автобусов марки «Мерседес Спринтер» рентабельность в целом крайне низкая.

Модель рентабельности пассажироперевозок с учетом управляемых и неуправляемых переменных будет иметь вид:

$$Y_i = Ef\{(x_1, x_2, \dots, x_k); (z_1, z_2, \dots, z_n)\} \quad (9)$$

Здесь x – управляемые, а z – неуправляемые факторы.

Проанализировав все x_i , пришли к выводу, что управляемыми являются только два x . Это x_1 – период времени, в который наиболее характерно меняется рентабельность пассажироперевозок и x_2 – списочное количество автобусов определенной марки на маршруте. Переменной x_1 прямо управлять нельзя, но можно выбирать определенный период времени и для него проводить расчет рентабельности.

В ходе исследований также было выяснено, что управляемая переменная x_2 будет входить в прогнозирующие модели в виде квадратичной зависимости.

Меняя управляемую переменную x_2 , можно максимизировать общую рентабельность пассажироперевозок в определенный промежуток времени (x_1): утренний час пик, день, вечер и т.п. И тем самым оптимизировать количество автобусов и их марок на определенных маршрутах предприятия-пассажироперевозчика.

Для каждой модели каждой марки автобусов с помощью разработанной автором итерационной процедуры поиска оптимального параметра регуляризации определили r_{opt} [3,4]. Зная r_{opt} для каждой модели $x_1, \dots, x_{30}$, были вычислены ридж-оценки коэффициентов.

Была проведена селекция прогнозирующих моделей с помощью критерия Ср Маллоуза. Это позволило определить вероятность выбора той или иной модели для расчета рентабельности. При селекции моделей использовалась бутстреп-процедура.

Были сделаны соответствующие расчеты для каждого определенного временного промежутка («утренний час пик», «день» и т.п.)

В результате для разных марок автобусов были получены модели вида (в качестве примера рассматривается период «Утро»):

$$\begin{aligned} Y_{\text{Э1}} &= 13.56 * \hat{Y}4E + 0.92 * \hat{Y}5E + 57.83 * \hat{Y}6E + 7.17 * \hat{Y}7E + 1.67 * \hat{Y}9E + \\ &+ 0.35 * \hat{Y}10E + 11.52 * \hat{Y}11E + 6.62 * \hat{Y}12E + 0.27 * \hat{Y}13E + 0.08 * \hat{Y}16E \\ Y_{\text{Б1}} &= 1.5 * \hat{Y}2B + 0.14 * \hat{Y}3B + 1.03 * \hat{Y}4B + 0.9 * \hat{Y}5B + 0.1 * \hat{Y}6B + 0.47 * \hat{Y}7B + \\ &+ 0.08 * \hat{Y}8B + 1.13 * \hat{Y}9B + 54.41 * \hat{Y}10B + 33.25 * \hat{Y}11B + 4.37 * \hat{Y}12B + 3.55 * \hat{Y}13B + \\ &+ 0.06 * \hat{Y}14B + 0.01 * \hat{Y}17B \\ Y_{\text{М1}} &= 71.01 * \hat{Y}1M + 15.87 * \hat{Y}2M + 5.93 * \hat{Y}3M + 6.44 * \hat{Y}4M + 0.33 * \hat{Y}6M + \\ &+ 0.14 * \hat{Y}9M + 0.24 * \hat{Y}10M + 0.02 * \hat{Y}12M + 0.01 * \hat{Y}13M + 0.01 * \hat{Y}16M \end{aligned} \quad (10)$$

Здесь в качестве коэффициентов перед переменными выступает вероятность выбора каждой определенной модели в процентах. Например, $\hat{Y}_{4B}$ – модель с четырьмя параметрами для марки автобусов «Богдан», $\hat{Y}_{2E}$ – модель с двумя параметрами для автобусов марки «Эталон» и т.д. и т.п.

Все x_i входят в модели нормированными, а x_2 в квадрате.

Все модели типа (10) являются вероятностными, т.к. истинная структура модели неизвестна. Модель с самым большим процентом выбора не используется в качестве результирующей, т.к. все модели входят в результирующую со своим процентным соотношением.

Проведенные исследования доказали, что критерий Ср Маллоуза работает, выбирая модели с наименьшей среднеквадратической ошибкой.

Для получения максимальной прибыли от пассажироперевозок анализируемого маршрута в каждый из исследуемых моментов времени, необходимо оптимизировать целевую функцию по формуле (7).

Меняя управляемую переменную x_2 , можно максимизировать Y_{max} – общую рентабельность пассажироперевозок в определенный промежуток времени (x_1): утренний час пик, день, вечер и т.п.

Таким образом, получили несколько систем уравнений, для соответствующих промежутков времени:

$$\begin{aligned} \text{Утро (1)} \quad & \begin{cases} Y_{1Э} = \alpha_{01Э} + \alpha_{11Э}x_{11Э} + \dots + \alpha_{n1Э}x_{n1Э} \\ Y_{1Б} = \alpha_{01Б} + \alpha_{11Б}x_{11Б} + \dots + \alpha_{n1Б}x_{n1Б} \\ Y_{1М} = \alpha_{01М} + \alpha_{11М}x_{11М} + \dots + \alpha_{n1М}x_{n1М} \\ \dots\dots\dots \end{cases} \\ \text{Суббота (5)} \quad & \begin{cases} Y_{5Э} = \alpha_{05Э} + \alpha_{15Э}x_{15Э} + \dots + \alpha_{n5Э}x_{n5Э} \\ Y_{5Б} = \alpha_{05Б} + \alpha_{15Б}x_{15Б} + \dots + \alpha_{n5Б}x_{n5Б} \\ Y_{5М} = \alpha_{05М} + \alpha_{15М}x_{15М} + \dots + \alpha_{n5М}x_{n5М} \end{cases} \end{aligned} \quad (11)$$

При этом $nЭ$, $nБ$, $nМ$ – это наборы разных переменных, так как значимость коэффициентов для различных моделей автобусов была разной для одинаковых x_i .

Количество автобусов определенной марки (x_2) также будет разным в различный период времени (x_1).

Определим управляемый x_2 через неуправляемые переменные по каждому периоду времени (x_1). Например, для периода времени 1 «Утро»:

$$\begin{cases} x_{21Э} = (Y_{1Э} - \alpha_{01Э} - \alpha_{31Э}x_{31Э} - \dots - \alpha_{n1Э}x_{n1Э}) / \alpha_{21Э} \\ x_{21Б} = (Y_{1Б} - \alpha_{01Б} - \alpha_{31Б}x_{31Б} - \dots - \alpha_{n1Б}x_{n1Б}) / \alpha_{21Б} \\ x_{21М} = (Y_{1М} - \alpha_{01М} - \alpha_{31М}x_{31М} - \dots - \alpha_{n1М}x_{n1М}) / \alpha_{21М} \end{cases} \quad (12)$$

В качестве ограничений будут выступать ограничения на управляемые переменные (например, количество автобусов определенной марки):

$$x_{2Э} \leq 15, \quad x_{2Б} \leq 12, \quad x_{2М} \leq 3.$$

$$\text{Тогда } Y_{\text{общ}} = \alpha_Y + x_{2Э} + x_{2Б} + x_{2М} \quad (13)$$

В коэффициент α_Y войдут все остальные иксы. Для получения коэффициентов для всех моделей x_i были стандартизированы.

Для того, чтобы получить рекомендации – в какое время, в каком количестве, автобусов каких марок необходимо вывести на маршрут для достижения максимальной прибыльности маршрута, необходимо решить все системы уравнений (11).

В рассматриваемой системе уравнений, в качестве каждого из x_i выступает количество автобусов определенной марки $x_{2Э}$, $x_{2Б}$, $x_{2М}$.

Для проведения оптимизации был использован метод линейного программирования. Так как в данной работе рассматриваются квази-стационарные процессы, было сделано допущение, что все функции, описывающие модели будут линейны, хотя бы на небольшом отрезке. Тогда в каждой итерации, меняя по одному автобусу каждой марки, была рассчитана функция рентабельности по формуле (13). Запоминался максимальный результат.

В результате решения системы (11) получили оптимальное количество автобусов всех имеющихся в наличии у предприятия марок на анализируемом маршруте 2 «А», для разного времени суток, позволяющее получить максимальную рентабельность пассажироперевозок в это время. Результаты вычислений представлены в таблице 1 и на рисунке 1.

Таблиця 1 – Результаты расчетов по определению оптимального количества автобусов на маршруте 2 «А» в разные временные периоды

Код периода	Название периода	Максимальная рентабельность, в %	Количество автобусов марки «Эталон», шт.	Количество автобусов марки «Богдан», шт.	Количество автобусов марки «Мерседес Спринтер», шт.
1	Утренний час пик	6.086	10	9	3
2	День	5.088	9	9	0
3	Вечерний час пик	6.088	10	9	3
4	Вечер	4.794	7	8	0
5	Суббота	5.165	9	9	2

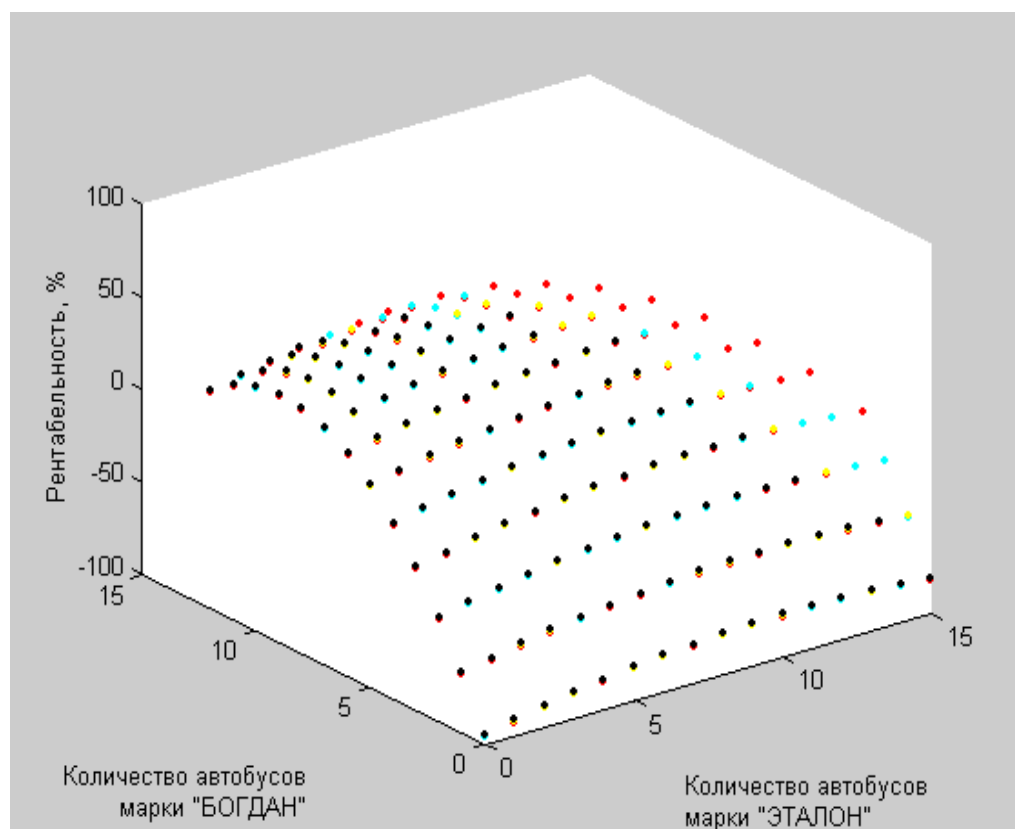


Рисунок 1 – График зависимости рентабельности пассажироперевозок от количества автобусов в период «День»

На рисунке количество автобусов марки «Мерседес Спринтер» определяется цветом точек: 0 автобусов марки «Мерседес» – красная точка; 1 автобус марки «Мерседес» – желтая точка; 2 автобуса марки «Мерседес» – синяя точка; 3 автобуса марки «Мерседес» – черная точка.

Были написаны программы на языке m – файлов для пакета MATLAB6p1, которые делают перебор всех данных (наблюдений) для определенного времени суток, находя и запоминая максимальную рентабельность (Y_i), при этом оптимизируя количество автобусов всех 3 марок. На графике рентабельность для наглядности умножена на 100.

Выводы и перспективы дальнейших исследований. Таким образом, доказано, что построенную модель рентабельности автотранспортного средства можно эффективно использовать для прогнозирования рентабельности пассажироперевозок в целом. Это также позволило оптимизировать численность автотранспортных средств на конкретных маршрутах анализируемого предприятия с точки зрения рентабельности.

Предприятию города были предложены конкретные рекомендации по требуемому количеству автобусов определенных марок в определенные временные периоды на различных маршрутах, а также

доказано, что автобус марки «Эталон» наиболее предпочтительно использовать на анализируемом маршруте. Данные рекомендации предприятие успешно использовало в процессе планирования объема пассажироперевозок на очередной период работы и приобретения новых автобусов. Это позволило сделать данный автобусный маршрут наиболее рентабельным, получить максимальную прибыль от пассажироперевозок. А также выявить ожидаемые изменения эксплуатационных показателей на маршруте. К ним относятся: продление маршрута и увеличение его протяженности; сокращение интервала движения в часы пик; повышение комфорта и снижение наполняемости автобусов в часы пик; повышение эксплуатационной скорости; изменение типа используемого автобуса; изменение средней дальности поездки пассажиров; увеличение объема перевозок на маршруте и другие показатели.

Располагая ожидаемыми изменениями эксплуатационных показателей, предприятие может разработать или уточнить эксплуатационный план пассажирских перевозок и по другим своим городским маршрутам. Полученные данные по объему перевозок можно в дальнейшем сопоставить с расчетом ожидаемой подвижности населения города. В соответствии с планом перевозок устанавливают потребное количество автобусов по типам, потребность в кадрах водителей и др.

Библиографический список использованной литературы

1. Дрейпер Н.Р. Прикладной регрессионный анализ / Н.Р. Дрейпер, Г. Смит; под ред. Н.Р. Дрейпера. — К.: Диалектика - Вильямс, 2007. — 912 с.
2. Hoerl A. Ridge regression: biased estimation for no orthogonal problems / A. Hoerl, R. Kennard // Technometrics. — 1970. — Vol. 12. — P. 55–67.
3. Потанина М.В. Построение модели оценки рентабельности автотранспортного средства методом смещенного оценивания параметров регрессионных уравнений / М.В. Потанина // Вестник СевНТУ. Экономика и финансы: сб. науч. тр. — Севастополь, 2008. — Вып. 92. — С. 166–171.
4. Потанина М.В. Оценка качества идентификации сложного объекта управления в условиях мультиколлинеарности / М.В. Потанина // Вестник СевНТУ. Сер. Автоматизация процессов управления: сб. науч. тр. — Севастополь, 2009. — Вып. 95. — С. 135–140.

Поступила в редакцию 01.03.2012 г.

Потаніна М.В. Підвищення рентабельності процесу міських автобусних пасажироперевезень

Розглянута методика оптимізації процесу міських автобусних пасажироперевезень з точки зору рентабельності автотранспортного засобу шляхом одночасного використання зміщеного оцінювання параметрів регресійних моделей і процедури селекції моделей оцінки рентабельності.

Ключові слова: Оптимізація, рентабельність, модель, селекція, регуляризація.

Potanina M.V. Increase of profitability of process of city bus passenger transportations due to his optimization

The method of optimization of process of city bus passenger transportations is in-process considered from point of profitability of vehicle by the simultaneous use of the displaced evaluation of parameters of regressive models and procedure of selection of models of estimation of profitability.

Keywords: Optimization, profitability, model, selection, regularization.