

Механика Транспорт Комуникации ^{Научно списание} ISSN 1312-3823 брой 1, 2011 г. статия № 0486 http://www.mtc-aj.com

ПРИЛОЖЕНИЕ НА МЕТОДА НА НАЙ-МАЛКИТЕ КВАДРАТИ ЗА РЕШАВАНЕ НА ОПТИМИЗАЦИОННИ ЗАДАЧИ В ТРАНСПОРТА С ИЗПОЛЗВАНЕ НА MS EXCEL SOLVER

Ставри Димитров

stavri_dimitrov@hotmail.com

Висше транспортно училище "Тодор Каблешков", катедра "Технология, организация и управление на транспорта", ул. "Гео Милев" 158, София 1574, БЪЛГАРИЯ

Ключови думи: приложение, метод на най-малките квадрати, ms excel, оптимизационни задачи, ms excel solver, единична линейна регресия, уравнение на регресия, регресионни коефициенти, транспорт

Резюме: Целта на настоящата публикация е да представи някои от възможностите на инструмента Microsoft® Excel Solver като средство за решаване на оптимизационни задачи в областта на транспорта. За постигане на целта, с емпирични данни от [3] е решен числен пример, в който по метода на най-малките квадрати, чрез използване на MS Solver, са изчислени оценките на неизвестните параметри на линеен еднофакторен регресионен модел, избран за аналитично описание на връзката между случайните променливи величини "скорост на движение на трамваен транспорт" и "средно разстояние между спирките".

1. Увод.

Съществуващото разнообразие от изчислителни програмни средства и специализирани статистически софтуерни пакети, предназначени основно за извършване на статистическа обработка на данни и статистически анализи като регресионен, корелационен, дисперсионен и др., предоставя на потребителите възможност за избор на подходящия за техните потребности софтуер. Една от възможните алтернативи за извършване на математически изчисления и статистически анализи е офис приложението *Microsoft*® *Excel* [4], а за решаване на оптимизационни задачи - вградения в *MS Excel* инструмент *MS Excel Solver*.

2. Предпоставки и начини за разрешаване на проблема.

Функционалност *MS Excel Solver* е от т.нар. *Add-In* инструменти, който след предварително инсталиране, е достъпен в програмата *MS Excel* и може да се използва за дефиниране и решаване на разнообразни по сложност и размерност оптимизационни задачи в различните научни области. Инструментът *MS Solver* дава възможност на потребителите, чрез разработен за това графичен потребителски интерфейс, ползвайки предварително въведените на работен лист (worksheet) в програмната среда на *MS Excel* входни данни, в диалогов режим да укажат търсената цел: минимум, максимум или точно определена стойност на зададения модел – линеен или нелинеен – и след

въвеждане на съпътстващите модела ограничения, лесно да дефинират и бързо да решат конкретна задача.

Настоящата публикация не представлява ръководство за работа със Solver. Представеният в публикацията материал е насочен не към описание на начина на работа на реализираните в *MS Excel Solver* изчислителни алгоритми, а по-скоро към тяхното практическо приложение при разрешаване на реални проблеми. Целта на публикацията е, чрез решаване на конкретен пример, да представи някои от възможностите за използване на *MS Solver* в областта на транспорта.

3. Решение на проучения проблем.

За постигане на поставената цел е решен и описан в публикацията следния числен пример, представящ примерно приложение на *MS Solver* в транспорта за изчисляване на коефициентите на уравнение на регресия: За показаните в табл. 1 средни стойности на разстоянията между спирките X_i (*m*) и скоростта на движение на трамваите Y_i (*km/h*), по метода на най-малките квадрати [1,2,3], чрез използване на *MS Excel Solver*, да се изчислят оценките на параметрите на неизвестен линеен еднофакторен регресионен модел.

X i	Yi	Xi	Yi	X i	Yi	
400	13,9	476	15,9	512	15,4	
402	14,2	478	15,5	516	16,6	
435	15,7	482	15,5	522	16,4	
445	15,1	483	15,5	526	16,4	
448	15,6	484	16,1	532	17,5	
450	15,4	492	15,7	533	17,2	
457	15,2	495	15,6	536	16,9	
457	15,7	497	16,6	539	15,4	
465	15,9	500	17,8	560	17,5	
469	15,3	506	15,2	562	16,5	

Таблица 1 – Емпирични данни

Нека за описание на връзката между двете променливи Y и X е избран линеен регресионен модел. Както знаем от [2], когато се работи с генерална съвкупност, неизвестният линеен регресионен модел има вида:

(1) $Y_i = \beta_0 + \beta_1 \cdot X_i + \varepsilon_i,$

където: β_0 и β_1 – неизвестни параметри на модела;

е_i – случаен компонент в модела.

За целите на разглеждания пример са обработени публикуваните в [3] данни от извършени 30 на брой наблюдения (табл. 1) върху стойностите на параметрите "средно разстояние между спирките" и "скорост на движение" в системата на градския обществен транспорт. В случаите като този, когато се работи с данни от направена извадка, за линейния регресионен модел, съгласно[2] можем да запишем:

(2)
$$Y_i = b_0 + b_1 \cdot X_i + e_i$$
,

в който: b_0 и b_1 – регресионни коефициенти, а e_i – остатъци.

Общият вид на линейното уравнение на регресия, използвано за описание на зависимостта на *Y* във функция от *X* е:

$$(3) \qquad \stackrel{\wedge}{Y_i} = b_0 + b_1 \cdot X_i$$

В примера скоростта на движение на превозните средства на трамвайния транспорт е зависимата (резултативната) променлива величина *Y*, а средното разстояние между спирките – независимата(факторната) променлива величина *X*.

Коефициентите b_0 и b_1 , в съответствие с [1], се явяват точкови оценки на неизвестните параметри b_0 и b_1 в регресионния модел (1). Стойностите на регресионните коефициенти b_0 и b_1 могат да се изчислят по метода на най-малките квадрати, при който се минимизират стойностите на отклоненията на емпиричните (наблюдаваните) точки Y_i от теоретичните (предвидените) стойности, лежащи на регресионната права, т.е. търси се минимум на сумата от квадратите на остатъците, обозначена в [2] като остатъчна девиация SSE:

(4)
$$SSE = \sum_{i=1}^{n} \left(Y_i - Y_i \right)^2 = \sum_{i=1}^{n} e_i^2 - \min$$

Съгласно метода на най-малките квадрати регресионната права се прокарва така, че да минава възможно най-близко до всички точки. Разположението и наклонът на правата в координатната система се определят от стойностите на регресионните коефициенти b_0 и b_1 .

След заместване на формула (3) в (4) получаваме:

(5)
$$SSE = \sum_{i=1}^{n} \left(Y_{i} - \hat{Y}_{i} \right)^{2} = \sum_{i=1}^{n} \left(Y_{i} - \left(b_{0} + b_{1} \cdot X_{i} \right) \right)^{2} = \sum_{i=1}^{n} \left(Y_{i} - b_{0} - b_{1} \cdot X_{i} \right)^{2} = \sum_{i=1}^{n} e_{i}^{2} - \min$$

Решението на задачата с използване на *MS Excel Solver* се намира след извършване на следната последователност от действия: Функционалност *Solver* се извиква от меню "*Tools*" (Инструменти) (фиг.1), от което се избира "*Solver*...".

🔀 Micros	oft Excel - A	pplication	of the L	east Squar	res Method fo	r Solvin	g Optimi	zation Ta	sks in th	e Transpo	rt Using M	S Excel S	olver
Eile 🔄	<u>E</u> dit <u>V</u> iew	Insert	F <u>o</u> rmat	Tools D	ata <u>W</u> indow	W <u>B</u> !	<u>H</u> elp						
1 🗋 💕	🔒 🔓	a 🗳	🍣 🖏	1 🔏 🖻	🖺 - 🝼 🖂	н (н	- f _x	έ 😫 Σ	z - Ag↓	X 🛍 🔺	🚯 50%	- 🕜]) i 🖬 🖬 🐿
Arial		• 16 •	BI	ū∣≣	≣ ≣ 🔤	\$ %	,€	00. 0. 0.♦ 00.	*	🔛 🗕 👌	• <u>A</u> •	, i ⊳	Security

Фиг. 1. Избор на меню "Инструменти"

Както беше обяснено по-горе *Solver* е *Add-In* инструмент. Това означава, че ако при отваряне на меню "*Tools*" *Solver* не се "вижда", т.е. не е добавен, същият следва да се добави. Това може да стане, след като от меню "*Tools*" се избере "*Add-Ins*…".

В отворилият се прозорец "Add-Ins..." (фиг. 2), с маркиране на "Solver Add-In" и поставяне на отметка, потребителят указва, че желае функционалността да бъде добавена. След като оптимизационния инструмент Solver е добавен, същият се пуска в действие, чрез избиране отново на меню "Tools"-> "Solver...". В резултат на това се показва прозореца "Solver Parameters" (Solver параметри) (фиг. 3).

Add-Ins	Solver Parameters	×
Add-Ins available:	Set Target Cell: Set Target Cell: Set Target Cell: Set Target Cell: Set	<u>S</u> olve Close
Solver Add-in	Subject to the Constraints:	Options
Solver Add-in Tool for optimization and equation solving	<u></u>	<u>R</u> eset All <u>H</u> elp

Фиг. 2.Прозорец "Add-Ins"

Фиг. 3.Прозорец "Solver параметри"

Тъй като в задачата се минимизира сумата от квадратите на остатъците, в прозореца "Solver Parameters" се избира "Min", а в полето "Set Target Cell:" като целева се указва клетката \$G\$34 (фиг. 4), в която Solver ще изчисли стойността на търсеният минимум. За успешно извършване на изчисленията е необходимо в целевата клетка предварително да е въведена формула (фиг. 4).

	A	В	C	D	E	F	G	н	I	J	К
1		T-5 0						-	T-S-mar O		
2		таолица и							тарлица з		
		,	.	τг.	Â				Регресионни коефициенти	Стойности	
3		· ·	A 2	11	Y_i	$Y_i - Y_i$	$(Y_i - Y_i)$			0.000	
4		1	400	13,9	0,0	13,900	193,210		60=	0,000	
5		2	402	14,2	0,0	14,200	201,640		b1 =	0,000	
6		3	435	15,7	0,0	15,700	246,490	⊢	J		
7		4	44)	15,1	0,0	15,100	228,010	⊢	Регресионно уравнение:		
*		<u> </u>	448	15,6	0,0	15,600	243,360	⊢	<u>^</u> , , ,		
9		6	450	15,4	0,0	15,400	237,160	⊢	$Y_i = b_0 + b_i X_i$		
10		7	457	15,2	0,0	15,200	231,040	⊢	1		
11		8	457	15,7	0,0	15,700	246,490	⊢			
12		9	465	15,9	0,0	15,900	252,810	⊢			
13		10	469	15,3	0,0	15,300	234,090	⊢			
14		11	476	15,9	0,0	15,900	252,810	⊢			
15		12	478	15,5	0,0	15,500	240,250	⊢			
16		13	482	15,5	0,0	15,500	240,250	⊢			
17		14	483	15,5	0,0	15,500	240,250	L			
18		15	484	16,1	0,0	16,100	259,210	L			
19		16	492	15,7	0,0	15,700	246,490	L			
20		17	495	15,6	0,0	15,600	243,360	L			
21		18	497	16,6	0,0	16,600	275,560	L			
22		19	500	17,8	0,0	17,800	316,840	L			
23		20	506	15,2	0,0	15,200	231,040	L			
24		21	512	15,4	0,0	15,400	237,160	L			
25		22	516	16,6	0,0	16,600	275,560	L			
26		23	522	16,4	0,0	16,400	268,960	L			
27		24	526	16,4	0,0	16,400	268,960	L			
28		25	532	17,5	0,0	17,500	306,250	L			
29		26	533	17,2	0,0	17,200	295,840	L			
30		27	536	16,9	0,0	16,900	285,610	L			
31		28	539	15,4	0,0	15,400	237,160	L			
32		29	560	17,5	0,0	17,500	306,250				
33		30	562	16,5	0,0	16,500	272,250				
34		$\sum_{i=1}^{n} \left(Y_i - Y_i \right)$	$\Big ^2 = \sum_{i=1}^n \Big _i$	$r_i = (b_0 + i)$	$(b_1 \cdot X_i)$	$=\sum_{i=1}^{n} e_i^2 =$	SUM(G4:G3	3)			
35		·	/ '¬\)	<i>.</i> ¬	SUM(numbe	er 1	; [number2];)		
36					љ (TT					

Фиг. 4.Начални стойности

Настройките по дефиниране на задачата приключват с посочване в полето "*By Changing Cells:*" (фиг. 5) на обхвата от клетки \$J\$4:\$J\$5 (табл. 3), съдържащи стойности, чрез промяна на които *Solver* ще намери минимум на функция (5), който ще върне като резултат във вече избраната целева клетка (\$G\$34).

Solver Parameters	×
Set Target Cell: \$G\$34 🔣	<u>S</u> olve
Equal To: O Max O Min O Value of: 0 By Changing Cells:	Close
\$J\$4:\$J\$5 Guess	
-Subject to the Constraints:	Options
Add	
⊆hange	
	<u>R</u> eset All
	Help

Фиг. 5.Прозорец "Solver параметри"

След като задачата е дефинирана, за начало на изчислителните процедури се натиска бутона "*Solve*". Намереният минимум (клетка G34) и стойностите на регресионните коефициенти (клетки J4 и J5), минимизиращи сумата от квадратите на остатъците, получени в резултат от извършените изчисления, са показани в табл. 4 и табл. 5 от работния лист (worksheet), представен на фиг. 6.

	A	В	C	D	E	F	G	H	1	J	к
1		ToEmus 4						-	ToEguus 5		
2		таолица 4					2	1	таолица э		
		<i>i</i> (7.	V.	Å.	л VV.	^^_ (vv.)		Регресионни коефициенти	Стойности	
5		· 1	400	12.0	145	0.547	0.201	┥┤	ka -	2 0 7 0	
4		1	400	13,9	14,5	-0,207	0,321	H	<u> </u>	7,970	
5		4	402	14,2	14,5	-0,299	0,090	₽	81 -	0,010	
6		<u>د</u>	430	15,7	15,0	0,000	0,442	⊢	D		
7		4	44)	15,1	15.2	-0,098	0,010	⊢	гегресионно уравнение:		
*			440	15,0	15.2	0,303	0,125	⊢	$V = \lambda_1 + \lambda_2 V = 707 +$.0016 V	
9		2	400	15,4	15,5	0,121	0,015	⊢	$I_i = 00 = 01 \cdot A_i = 1, y_i = 1$	0,010.2	
10		, ,	457	15.2	15,4	-0,195	0,037	⊢	1		
11		<u> </u>	451	15,7	15,4	0,307	0,094	⊢			
12		9	400	15.9	15,5	0,377	0,142	⊢			
13		10	409	15,5	15.0	-0,288	0,083	⊢			
14		11	470	15,9	15.7	0,199	0,039	⊢			
15		12	478	15,5	15.0	-0,234	0,000	⊢			
16		13	482	15,5	15.0	-0,299	0,089	⊢			
17		14	483	10,0	15.0	-0,312	0,099	⊢			
18		15	484	10,1	15,8	0,209	0,072	⊢			
19		10	492	15,7	16,0	-0,201	0,068	⊢			
20		17	495	15,6	16,0	-0,410	0,168	⊢			
21		18	497	10,0	10,0	0,007	0,311	⊢			
22		19	200	17,8	16,1	1,709	2,920	⊢			
23		20	510	15,2	16,2	-0,989	0,978	⊢			
24		21	512	15,4	10,3	-0,880	0,785	⊢			
25		22	500	10,0	10,4	0,249	0,002	⊢			
26		23	524	16,4	10,4	-0,049	0,002	⊢			
27		24	520	10,4	16,5	-0,114	0,013	⊢			
28		20	532	17,5	10,0	0,889	0,790	⊢			
29		20	233	17,2	16,0	0,273	0,328	⊢			
30		21	530	16,9	10,7	0,224	0,000	⊢			
31		28	239	10,4	10,7	-1,320	1,700	⊢			
32		29	260	17,5	17,1	0,434	0,189	⊢			
33		30	262	10,0	17,1	-0,298	0,358	⊢			
		<u>n</u> (., ^) ² _~~(.	. ()	_ <u>n</u> _2_	10,490	I			
34		$\sum_{i=1}^{N} Y_i - Y_i $) "것[]	(i-160 +i	$b_1 \cdot X_i$	$\sum_{i=1}^{n} e_i = $					
35		,-1(] I	/ '¬\	1	, /	2-1		_			
36				<u>т</u> (D			1			

Фиг. 6.Резултати от изчисленията

Индикация на успешното приключване на изчисленията е изведеното в отвореният прозорец "Solver Results" (фиг. 7) текстовото съобщения указващо, че е намерено оптимално решение, удовлетворяващо всички ограничения. За съхраняване

на решението се избира предложеният от Solver избор "Keep Solver Solution" и се натиска бутона "OK".

Solver Results			×
Solver has converged to the current so constraints are satisfied.	olution. All	<u>R</u> eports	
		Answer Sensitivity Limits	*
OK Cancel	Save Scenario	·	Help

Фиг. 7.Прозорец "Solver Results"

За онагледяване на крайните резултати, на построената на фиг. 8 диаграма са показани емпиричните (наблюдаваните) стойности и регресионната права, на която лежат теоретичните (предвидените) стойности, описваща се от полученото след заместване на изчислените коефициенти b_0 и b_1 линейно еднофакторно уравнение на регресия, имащо следния запис:



4. Резултати и дискусия.

Получените от решението на задачата, чрез използване на *MS Excel Solver* резултати показват, че:

1. Евентуалното увеличаване/намаляване на средното разстояние между спирките с 1 m ще доведе до увеличаване/намаляване на скоростта на движение на превозните средства на трамвайния транспорт средно с 0,016 km/h.

2. При нулево разстояние между спирките скоростта на движение ще е равна на 7,97 *km/h*, което означава, че наличието на свободен член прави подбраният линеен регресионен модел неподходящ за моделиране на връзката между скоростта на движение на трамваите и средното разстояние между спирките.

3. Потвърждение на горното е и факта, че построената регресионна права(фиг. 8) не се вписва добре в емпиричните точки, от което може да се направи предположение, че връзката между променливите *Y* и *X* не е силна, или ако е силна, избраното линейно регресионно уравнение не я описва адекватно.

4. За извършване на по-точна апроксимация вероятно е необходимо да се извършат повече на брой наблюдения за значенията на променливите *Y* и *X*, както и да се избере и тества друг вид регресионен модел - нелинеен еднофакторен или линеен (или нелинеен) многофакторен модел, тъй като освен от разстоянието между спирките, е възможно скоростта на движение да се влияе и о т други неотчетени от модела фактори.

5. Заключение.

Резултатите от примерното приложение на функционалност *MS Excel Solver* дават основание да се направят следните основни изводи:

1. Практическото приложение на съществуващите софтуерни програмни средства – специализирани или с общо предназначение, заедно с вградените в тях функционалности – представляват алтернатива на познатите до момента ръчни способи за извършване изчисления и анализи в научните изследвания.

2. Използването на електронните таблици *MS Excel* и на добавъчния към него инструмент *MS Excel Solver*, като средство за извършване на математически изчисления и решаване на оптимизационни задачи, се характеризира със следните предимства и особености:

- тъй като *MS Excel Solver* е допълнителен компонент към едно от найширикоразпространените и използвани от потребителите офис приложения *MS Excel*, то същият е леснодостъпен;

- удобно и лесно се работи със *Solver*, което го прави подходящ не само за използване от инженери, математици, икономисти и изследователи в провежданите от тях изследвания, но и от студенти за учебни цели;

- инструментът *Solver* позволява със сравнително малко усилия да се дефинират и решават разнообразни по вид и сложност задачи;

- в зависимост от начално присвоените стойности в клетки J4 и J5 (в случая клетките в табл. 3, съдържащи стойностите на регресионните коефициенти) и предварително зададения брой на извършваните от *Solver* итерации, е възможно изчислените коефициенти незначително да се различават по стойност от същите, но получени в резултат от използване на други изчислителни програми.

Разбира се, чрез описания в настоящата публикация пример не се изчерпват възможностите за използване на *Solver* и съществуващите методи за изчисляване на стойностите на регресионните коефициенти.

ЛИТЕРАТУРА

- [1] Качаунов, Т. Т., "Моделиране и оптимизация на транспортните процеси", второ преработено издание, Печатница при ВТУ "Тодор Каблешков", София, 2005 г.
- [2] Манов, А., "Статистика със SPSS", второ издание, Издателство "Тракия М", София, 2001 г.
- [3] Рихтер, К. Ю., Фишер, П., Шнейдер, Г., "Статистические методы в транспортных исследованиях", превод от немски, Москва, "Транспорт", 1982 г.
- [4] Microsoft Excel 2003 Help, Microsoft Corporation