Дополнительные возможности анализа данных в MS Excel

1. Аппроксимация экспериментальных данных. Линии тренда

На практике часто приходится сталкиваться с задачей о сглаживании экспериментальной зависимости или задачей аппроксимации.
Аппроксимация (от лат. Approximo - приближение) - замена одних математических объектов (например, чисел или функций) другими, более простыми и в том или ином смысле близкими к исходным (напр., кривых линий близкими к ним ломаными).
Аппроксимация - приближенное решение сложной функции с помощью более простых, резко ускоряет и упрощает решение задач. В экономике целью аппроксимации часто является укрупнение характеристик моделируемых экономических объектов.

Аппроксимацией называется процесс подбора эмпирической формулы для установленной из опыта функциональной зависимости y = f (x). Эмпирические формулы служат для аналитического представления экспериментальных данных.

В простейшем случае задача аппроксимации экспериментальных данных выглядит следующим образом.

Пусть какие-то данные, полученные практическим путем (в ходе эксперимента или наблюдения), можно представить парами чисел (x, y). Зависимость между ними отражает таблица 2.

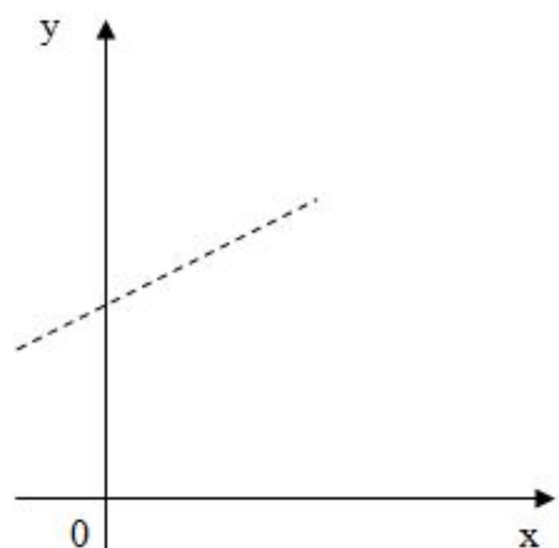
Таблица 2 - Зависимость экспериментальных X и Y

X	X_1	X_2	...	X_n
Y	Y_1	Y_2	...	Y_n

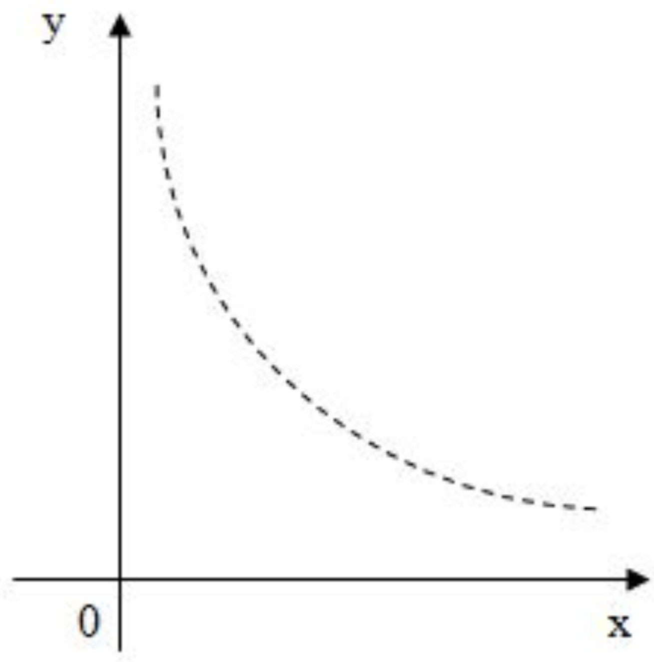
На основе этих данных нужно подобрать функцию y = f (x), которая наилучшим образом сглаживала бы экспериментальную зависимость между переменными и по возможности точно отражала общую тенденцию зависимости между x и y, исключая погрешности измерений и случайные отклонения. Это значит, что отклонения в каком-то значении $y_i - y_i(x_i)$ бы минимальными.

Выяснить вид функции можно либо из теоретических соображений, или анализируя расположение точек (x n; yn) на координатной плоскости.

Например, пусть точки расположены так:



Учитывая то, что практические данные получены с некоторой погрешностью, обусловленной неточностью измерений, необходимостью округления результатов и т. п., естественно предположить, что здесь имеет место линейная зависимость $y = ax + b$. Чтобы функция приняла конкретный вид, необходимо каким-то образом вычислить a и b.



Расположение экспериментальных точек может иметь самый разный вид, и каждому соответствует конкретный тип функции.

Построение эмпирической функции сводится к вычислению параметров, входящих в нее, так чтобы из всех функций такого вида выбрать ту, которая лучше других описывает зависимость между величинами, которые изучаются. То есть сумма квадратов разницы между табличными значениями функции в некоторых точках и значениями, вычисленными по полученной формуле, должна быть минимальной.

В MS Excel аппроксимация экспериментальных данных осуществляется путем построения их графика (х - абстрактные величины) или **точечного графика (х - имеет конкретные значения) с последующим подбором соответствующей аппроксимирующей функции (линии тренда).**

Возможны следующие варианты функций:

Линейная - $y = ax + b$. Обычно применяется в простейших случаях, когда экспериментальные данные растут или убывают с постоянной скоростью.

Полиномиальная — $y=a_0 + a_1x^2+a_2x^2 +... + a_nx^n$, где полиномы до шестого порядка включительно (n <= 6), a_i — константы. Используется для описания экспериментальных данных, попеременно возрастающих и убывающих. Степень полинома определяется количеством экстремумов (максимумов или минимумов) кривой. Полином второй степени может описать только один максимум или минимум, полином третьей степени может иметь один или два экстремума, четвертой степени - не более трех экстремумов и т. .

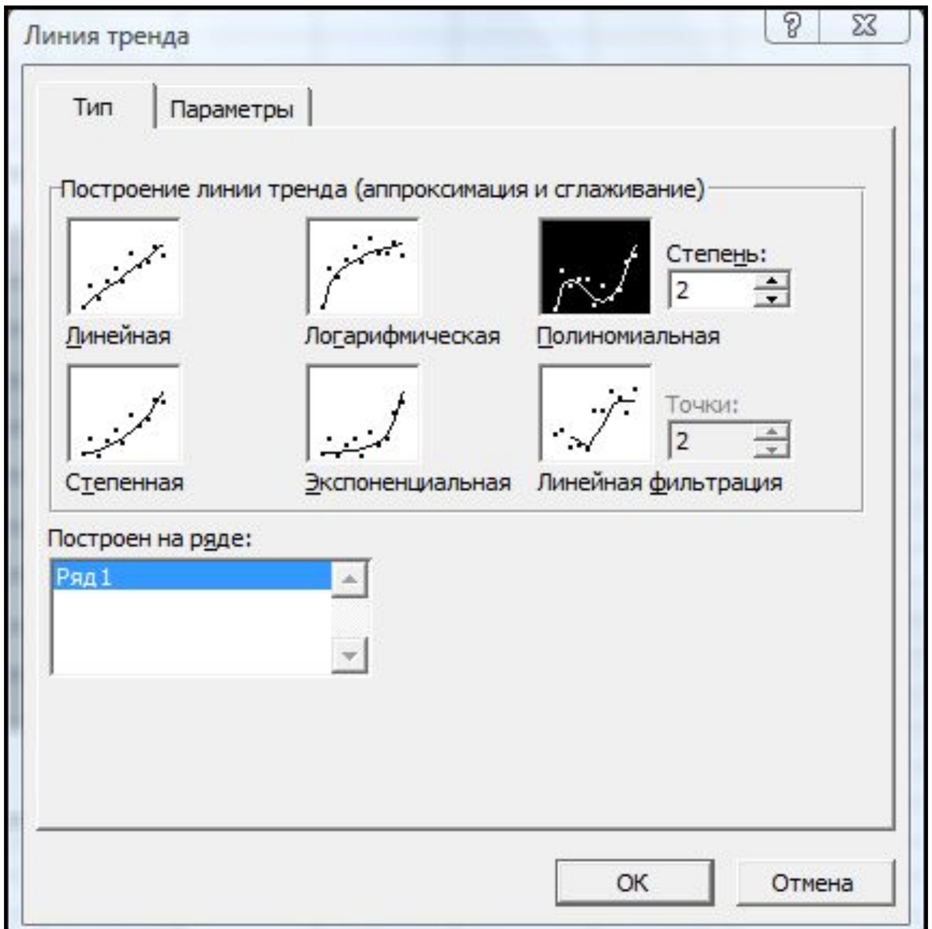
Логарифмическая - $y = a\ln x + b$, где a и b - константы, ln - функция натурального логарифма. Функция применяется для описания экспериментальных данных, которые сначала быстро растут или убывают, а затем постепенно стабилизируются.

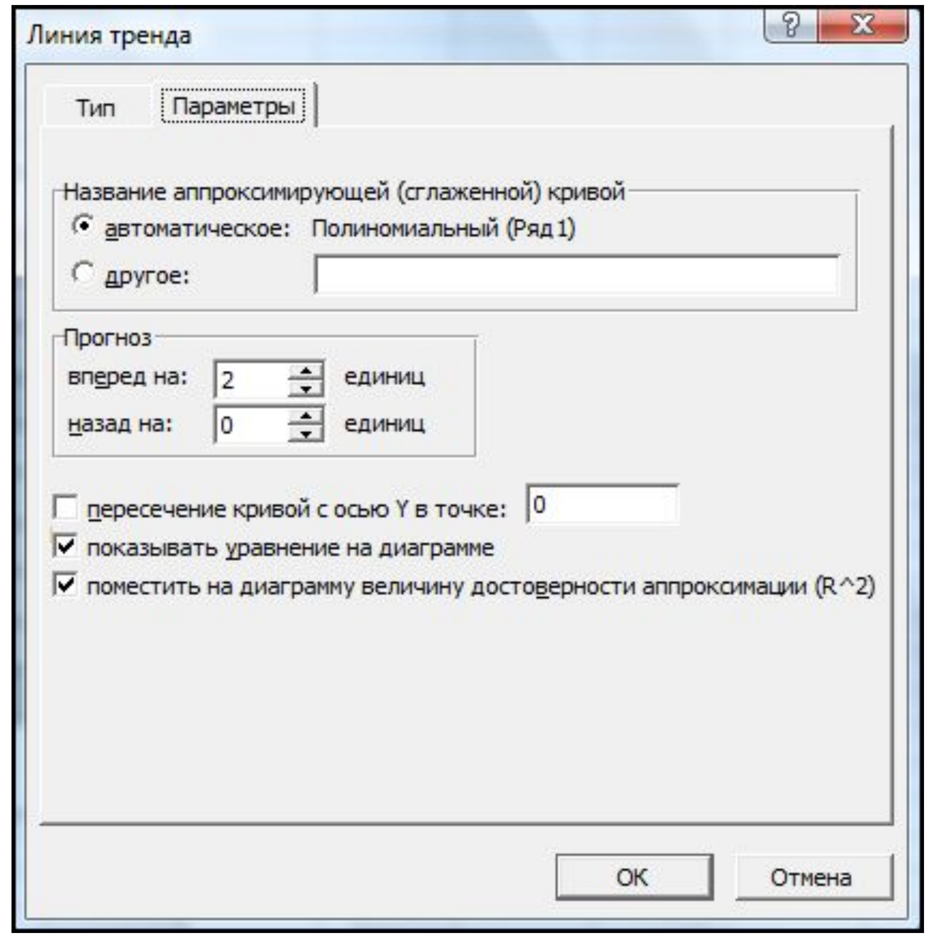
Степенная — $y = bx^a$, де *a* і *b* — константы. Аппроксимация степенной функцией используется для экспериментальных данных, со скоростью роста, которая постоянно увеличивается (или убывает). Данные не должны иметь нулевых или отрицательных значений.

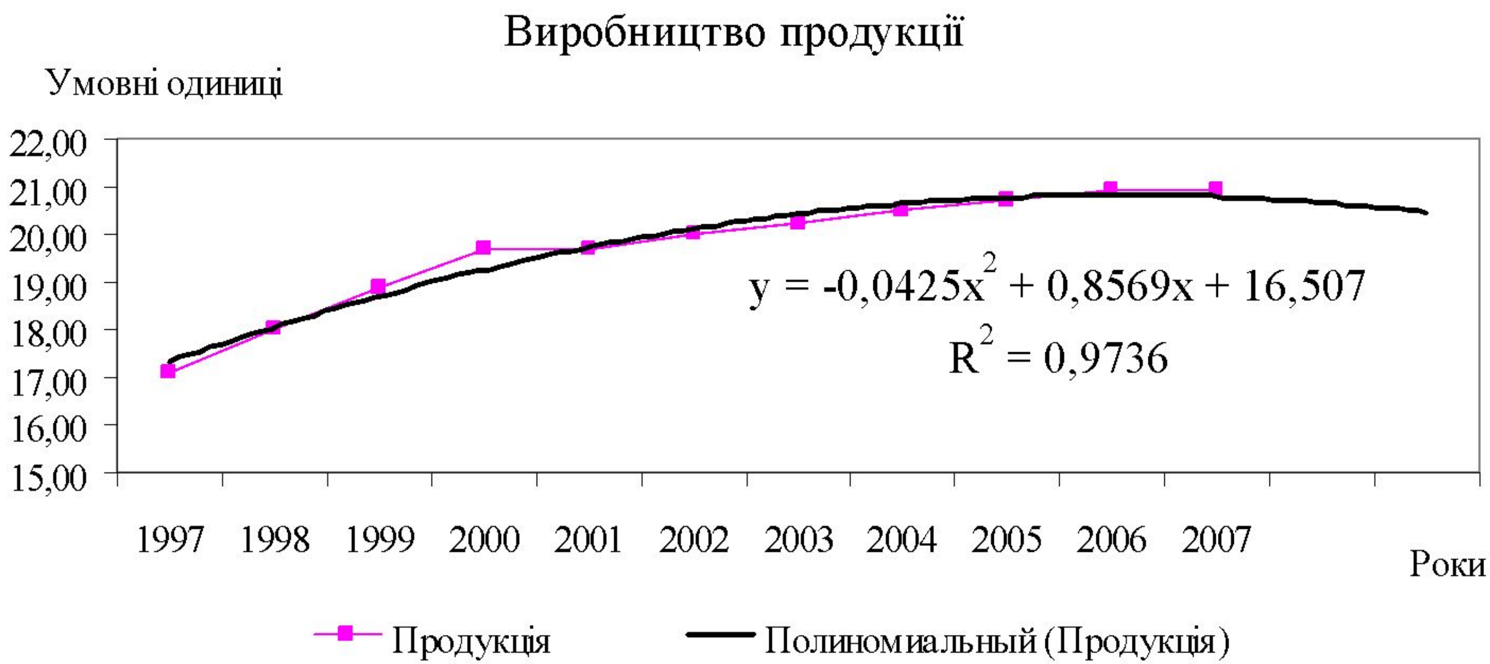
Экспоненциальная — $y= be^{ax}$, где *a* и *b* — константы, *e* — основание натурального логарифма. Применяется для описания экспериментальных данных, которые быстро растут или убывают, а затем постепенно стабилизируются. Часто ее использование следует из теоретических соображений.

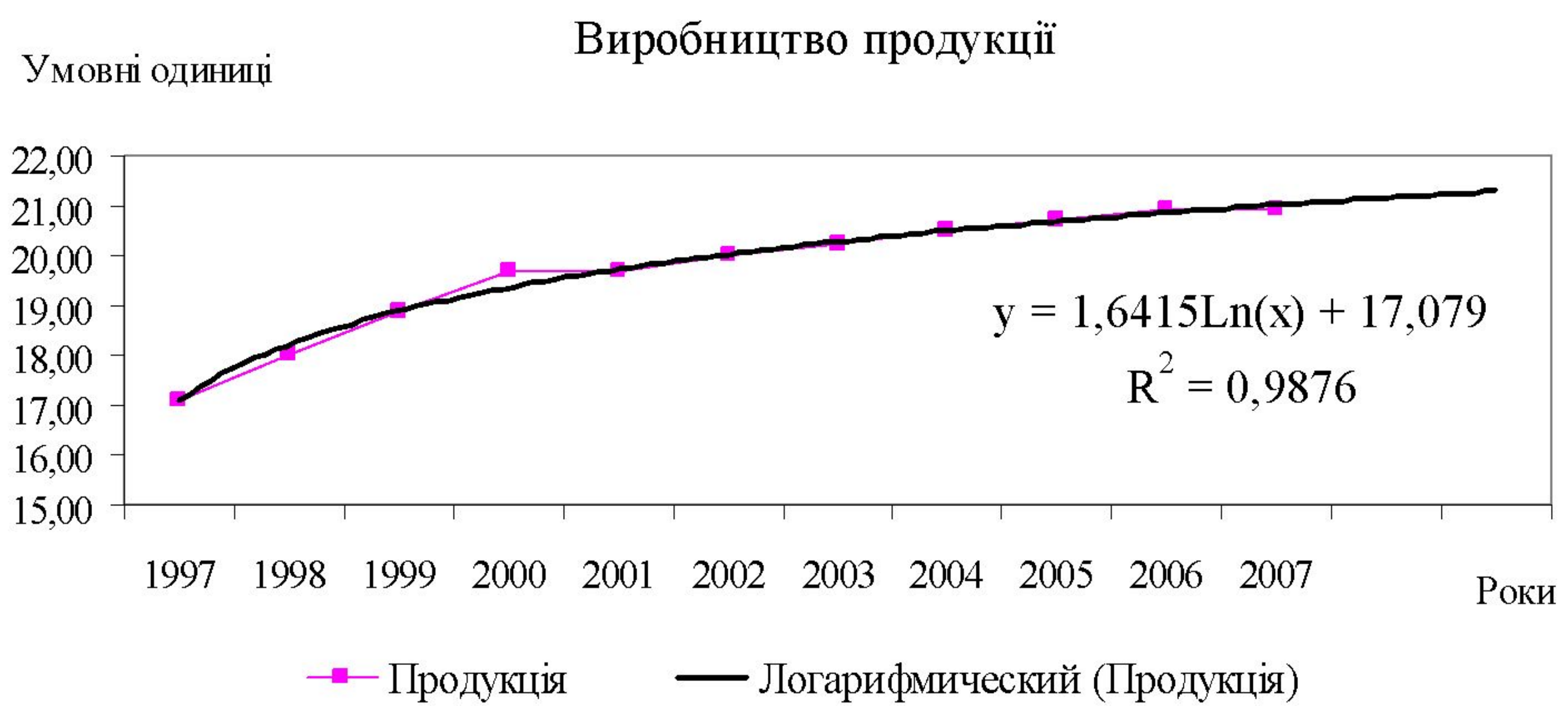
Степень близости аппроксимации экспериментальных данных выбранной функции оценивается коэффициентом детерминации (R^2). Таким образом, если есть несколько подходящих вариантов типов аппроксимирующих функций, можно выбрать функцию с большим коэффициентом детерминации (стремящимся к 1).

Для осуществления аппроксимации на диаграмме экспериментальных данных необходимо щелчком правой кнопки мыши вызвать контекстное меню и выбрать пункт **Добавить линию тренда**. В диалоговом окне **Линия тренда** на вкладке **Тип** выбирается вид аппроксимирующей функции, а на вкладке **Параметры** задаются дополнительные параметры, влияющие на отображение аппроксимирующей кривой.






Виробництво продукції
Умовні одиниці
22,00
21,00
20,00
19,00
18,00
17,00
16,00
15,00
y = -0,0425x^2 + 0,8569x + 16,507
R^2 = 0,9736
1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Роки
Продукція
Полиномиальный (Продукція)


Виробництво продукції
Умовні одиниці
22,00
21,00
20,00
19,00
18,00
17,00
16,00
15,00
y = 1,6415Ln(x) + 17,079
R^2 = 0,9876
1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007
Роки
Продукція
Логарифмический (Продукція)

2. Поиск решения задач линейного программирования (ЗЛП)

Теория ЗЛП

Оптимизация — в математике, информатике и исследовании операций задача нахождения экстремума (минимума или максимума) целевой функции в некоторой области конечномерного векторного пространства, ограниченной набором линейных и/или нелинейных равенств и/или неравенств.

Теорию и методы решения задачи оптимизации изучает ***математическое программирование***.

Математическое программирование - это область математики, разрабатывающая теорию, численные методы решения многомерных задач с ограничениями. В отличие от классической математики, математическое программирование занимается математическими методами решения задач нахождения наилучших вариантов из всех возможных.

В процессе проектирования ставится обычно задача определения наилучших, в некотором смысле, структуры или значений параметров объектов. Такая задача называется **оптимизационной**.

Для того, чтобы корректно поставить задачу оптимизации, необходимо задать:

- *допустимое множество*;
- *целевую функцию*;
- *критерий поиска* (max или min).

Задачи оптимизации, в которых целевая функция и ограничения являются линейными функциями, разрешаются так называемыми методами *линейного программирования*.

Краткая история вопроса

Математические исследования отдельных **экономических** проблем, математическая формализация числового материала проводилась ещё в XIX веке.

В 1939 году Л. В. Канторович опубликовал работу «**Математические методы организации и планирования производства**», таким образом были заложены основы линейного программирования.

Изучение подобных задач привело к созданию новой научной дисциплины **линейного программирования** и открыло новый этап в развитии **экономико-математических методов.**

Термин «программирование» нужно понимать в смысле «планирования» (один из переводов англ. *programming*). Он был предложен в середине 1940-х годов Д. Данцигом, одним из основателей линейного программирования, **ещё до того, как компьютеры были использованы для решения линейных задач оптимизации.**

Практика решения ЗЛП

1 часть - Построение математических моделей линейного программирования

.

Математические модели линейного программирования строятся на основе содержательной постановки экономической задачи.

Пример.

Типичная **задача распределения ресурсов** ставится следующим образом.

Пусть фирма выпускает продукцию четырех типов **Продукт1, Продукт2, Продукт3, Продукт4**, для изготовления которой требуются **ресурсы трех видов: трудовые, сырье, финансы.** Количество ресурса каждого вида, необходимое для выпуска единицы продукции данного типа, называется **нормой расхода**.

Исходные данные.

Норма расхода, а также прибыль, получаемая от реализации единицы каждого типа продукции, приведены в таблице 1, там же приведен объем ресурса, которым можно располагать. **Требуется определить**, в каком **количестве надо выпускать** продукцию каждого типа, **чтобы суммарная прибыль была максимальной**.

Таблица 1 – Исходные данные задачи линейного программирования

Ресурс	Про-дукт1	Про-дукт2	Продукт3	Продукт4	Наличие
Трудовые	1	1	1	1	16
Сырье	6	5	4	3	110
Финансы	4	6	10	13	100
Прибыль	60	70	120	130	

Составление математической модели начинают с выбора переменных задачи. В большинстве случаев (для ресурсных задач в частности) переменные – неизвестные количества. После выбора переменных, исходя из содержательной формулировки задачи, последовательно составляют линейные ограничения, которые эти переменные должны удовлетворять.

При этом нужно следить, чтобы в модель были введены все ограничительные условия и в то же время не было ни одной лишней или записанной в более жесткой форме, чем требуется по условиям задачи.

Следующим шагом является составление целевой функции, которая в математической форме отражает заданный в условиях задачи критерий оптимизации и должна быть линейной.

Заметим, что в некоторых моделях удобнее целевую функцию строить сразу после выбора переменных задачи, т.е. порядок построения модели не является жестким и может изменяться.

После построения модель, если это возможно, упрощают.

Составим математическую модель, для чего введем следующие обозначения:

x_j- количество выпускаемой продукции j-го типа $j=1,2,3,4$;

b_i- количество располагаемого ресурса i-го вида $i=1,2,3$;

a_{ij}- норма расхода i-го ресурса для выпуска единицы продукции j-го типа;

c_j- прибыль, получаемая от реализации единицы продукции j-го типа.

Из таблицы 1 видно, что для выпуска единицы Продукта1 требуется 6 единиц сырья, значит, для выпуска всей продукции первого типа требуется **$6x_1$ единиц сырья**, где **x_1- количество** выпускаемой продукции **Продукт1**. С учетом того, что для других видов продукции зависимости будут аналогичны, ограничение по сырью будет иметь вид:

$6 \cdot x1+5 \cdot x2+4 \cdot x3+3 \cdot x4 \leq 110.$

В этом ограничении левая часть равна величине требуемого ресурса, а правая показывает количество имеющегося ресурса.

Аналогично можно составить ограничения для остальных ресурсов и написать зависимость для целевой функции.

Математическая модель задачи выглядит следующим образом.

Целевая функция имеет вид:

$60 \cdot x_1 + 70 \cdot x_2 + 120 \cdot x_3 + 130 \cdot x_4 \rightarrow max$

Ограничения имеют вид:

$x_1 + x_2 + x_3 + x_4 \le 16$

$6 \cdot x_1 + 5 \cdot x_2 + 4 \cdot x_3 + 3 \cdot x_4 \le 110$

$4 \cdot x_1 + 6 \cdot x_2 + 10 \cdot x_3 + 13 \cdot x_4 \le 100$

$x_j \ge 0;\ j=1,4$.

2 часть – Решение задачи линейного программирования с помощью табличного процессор MS Excel

Ввод условий задачи состоит из следующих основных шагов:

1) Создание формы для ввода условий задачи.

2) Ввод исходных данных (коэффициентов математической модели).

3) Ввод целевой функции, ограничений и граничных условий.

Форма ввода ЗЛП

	A	B	C	D	E	F	G	H	I
1			Переменные						
2	**Продукция**	Продукт 1	Продукт 2	Продукт 3	Продукт 4				
3	Количество								
4	Нижняя гр.								
5	Верхняя гр.					ЦФ	направл.		
6	Коэф. В ЦФ								
7			Ограничения задачи						
8	**Ресурс**	Продукт 1	Продукт 2	Продукт 3	Продукт 4	**Левая часть**	**Знак**	**Правая часть**	
9	Трудовые								
10	Сырье								
11	Финансы								
12									

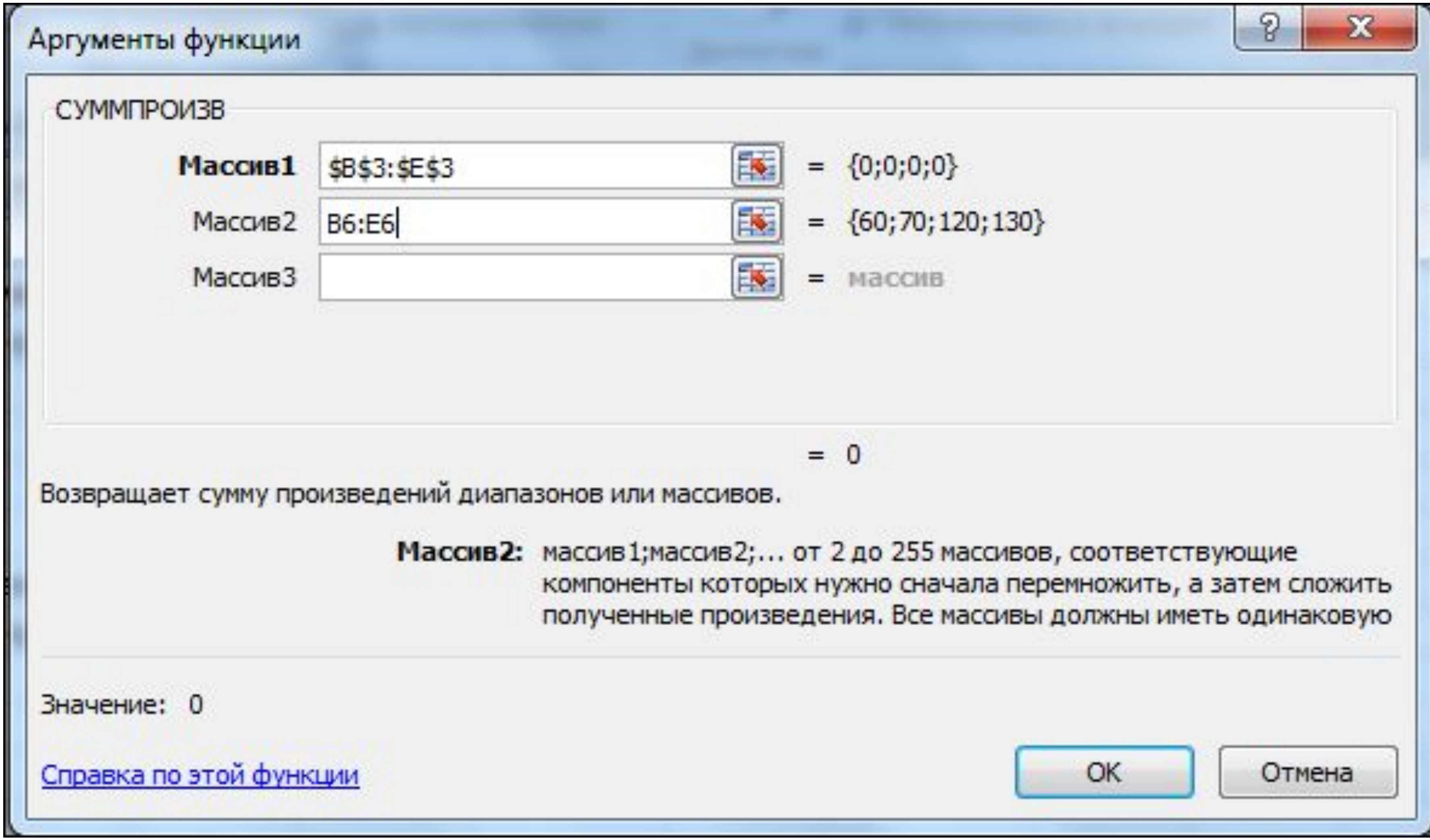
Исходные данные

	A	B	C	D	E	F	G	H
1		Переменные						
2	**Продукция**	Продукт 1	Продукт 2	Продукт 3	Продукт 4			
3	Количество	0	0	0	0			
4	Нижняя гр.	0	0	0	0			
5	Верхняя гр.					ЦФ	направл.	
6	Коэф. В ЦФ	60	70	120	130			
7		Ограничения задачи						
8	**Ресурс**	Продукт 1	Продукт 2	Продукт 3	Продукт 4	**Левая часть**	**Знак**	**Правая часть**
9	Трудовые	1	1	1	1		<=	16
10	Сырье	6	5	4	3		<=	110
11	Финансы	4	6	10	13		<=	100
12								

Ввод уравнений целевой функции и граничных условий

	A	B	C	D	E	F	G	H
1		Переменные						
2	**Продукция**	Продукт 1	Продукт 2	Продукт 3	Продукт 4			
3	Количество	0	0	0	0			
4	Нижняя гр.	0	0	0	0			
5	Верхняя гр.					ЦФ	направл.	
6	Коэф. В ЦФ	60	70	120	130	0		
7		Ограничения задачи						
8	**Ресурс**	Продукт 1	Продукт 2	Продукт 3	Продукт 4	**Левая часть**	**Знак**	**Правая часть**
9	Трудовые	1	1	1	1	0	<=	16
10	Сырье	6	5	4	3	0	<=	110
11	Финансы	4	6	10	13	0	<=	100
12								

	A	B	C	D	E	F	G	H
1	Переменные							
2	**Продукция**	Продукт 1	Продукт 2	Продукт 3	Продукт 4			
3	Количество	0	0	0	0			
4	Нижняя гр.	0	0	0	0			
5	Верхняя гр.					ЦФ	направл.	
6	Коэф. В ЦФ	60	70	120	130	**=СУММПРОИЗВ($B$3:$E$3;B6:E6)**		
7	Ограничения задачи							
8	**Ресурс**	Продукт 1	Продукт 2	Продукт 3	Продукт 4	**Левая часть**	**Знак**	**Правая часть**
9	Трудовые	1	1	1	1	=СУММПРОИЗВ($B$3:$E$3;B9:E9)	<=	16
10	Сырье	6	5	4	3	=СУММПРОИЗВ($B$3:$E$3;B10:E10)	<=	110
11	Финансы	4	6	10	13	=СУММПРОИЗВ($B$3:$E$3;B11:E11)	<=	100


Аргументы функции
СУММПРОИЗВ
Массив1
$B$3:$E$3
= {0;0;0;0}
Массив2
B6:E6
= {60;70;120;130}
Массив3
= массив
= 0
Возвращает сумму произведений диапазонов или массивов.
Массив2: массив1;массив2;... от 2 до 255 массивов, соответствующие компоненты которых нужно сначала перемножить, а затем сложить полученные произведения. Все массивы должны иметь одинаковую
Значение: 0
Справка по этой функции
ОК
Отмена

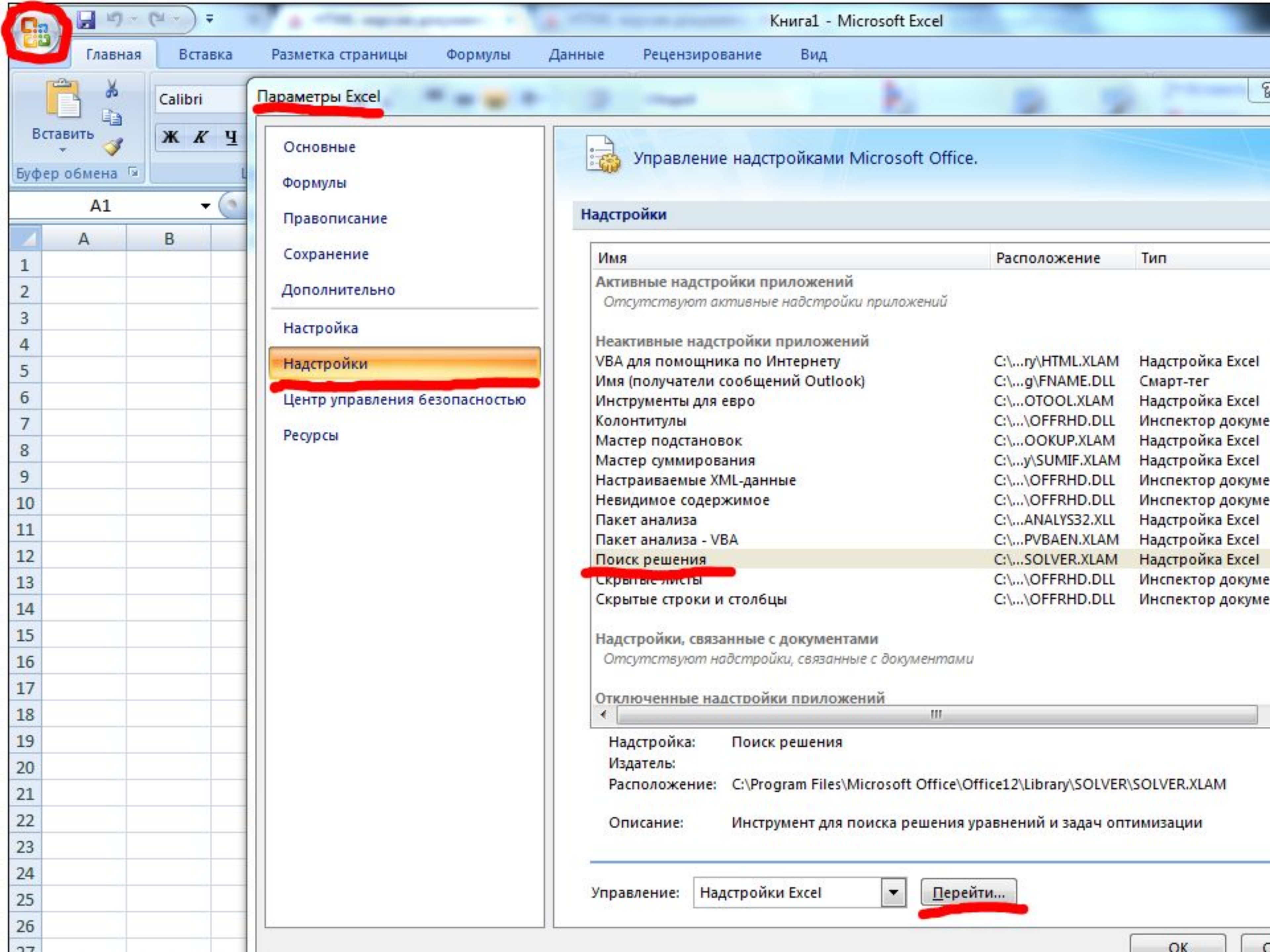
Дальнейшее решение задачи проводится с использованием надстройки <u>Поиск решения</u> табличного процессора MS Excel.

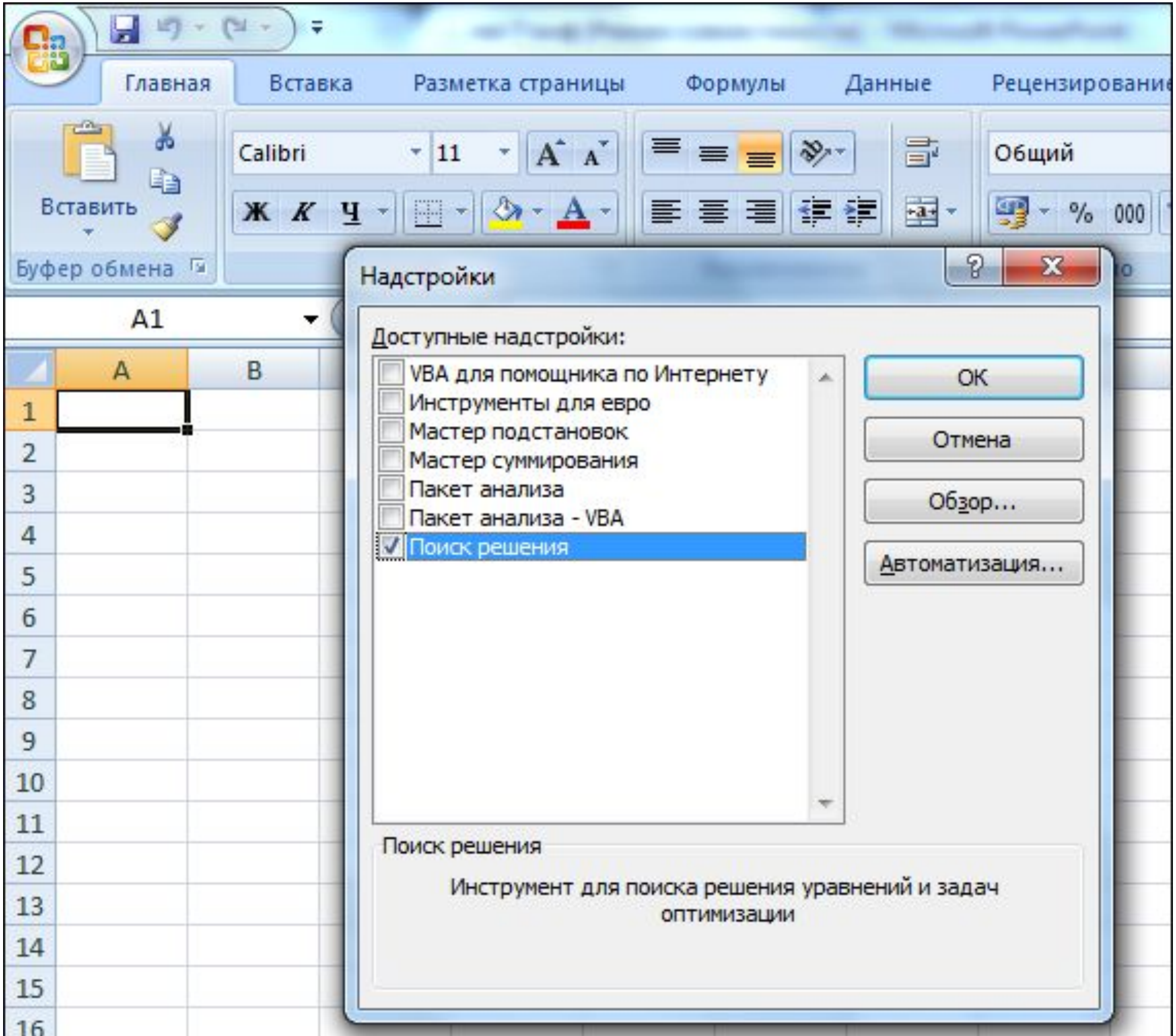
<u>Поиск решения</u> является надстройкой MS Excel, инструментом для поиска решения уравнений и задач оптимизации.

Данная надстройка, в случае отсутствия в пунктах меню MS Excel 2003 и более ранних версий, может быть добавлена с помощью выбора пунктов Сервис – Надстройки – Поиск решения.

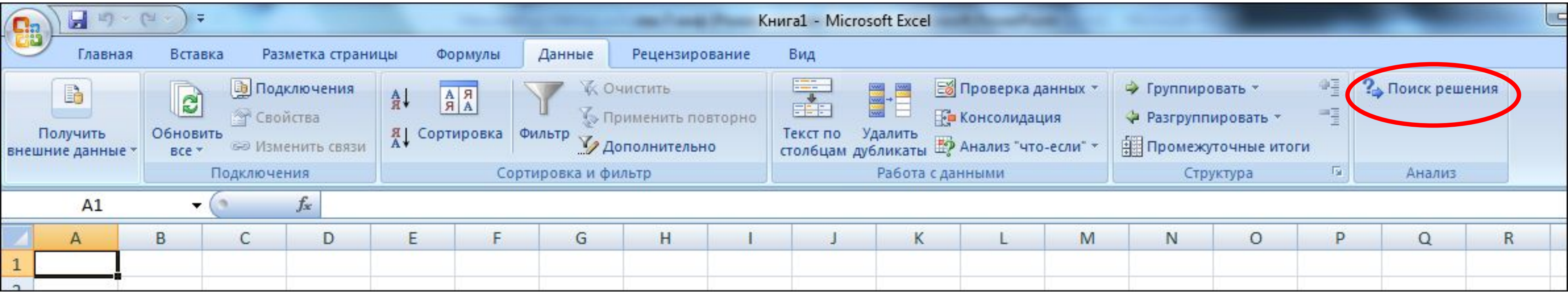
После этого будет добавлен подпункт Поиск решения в меню Сервис.

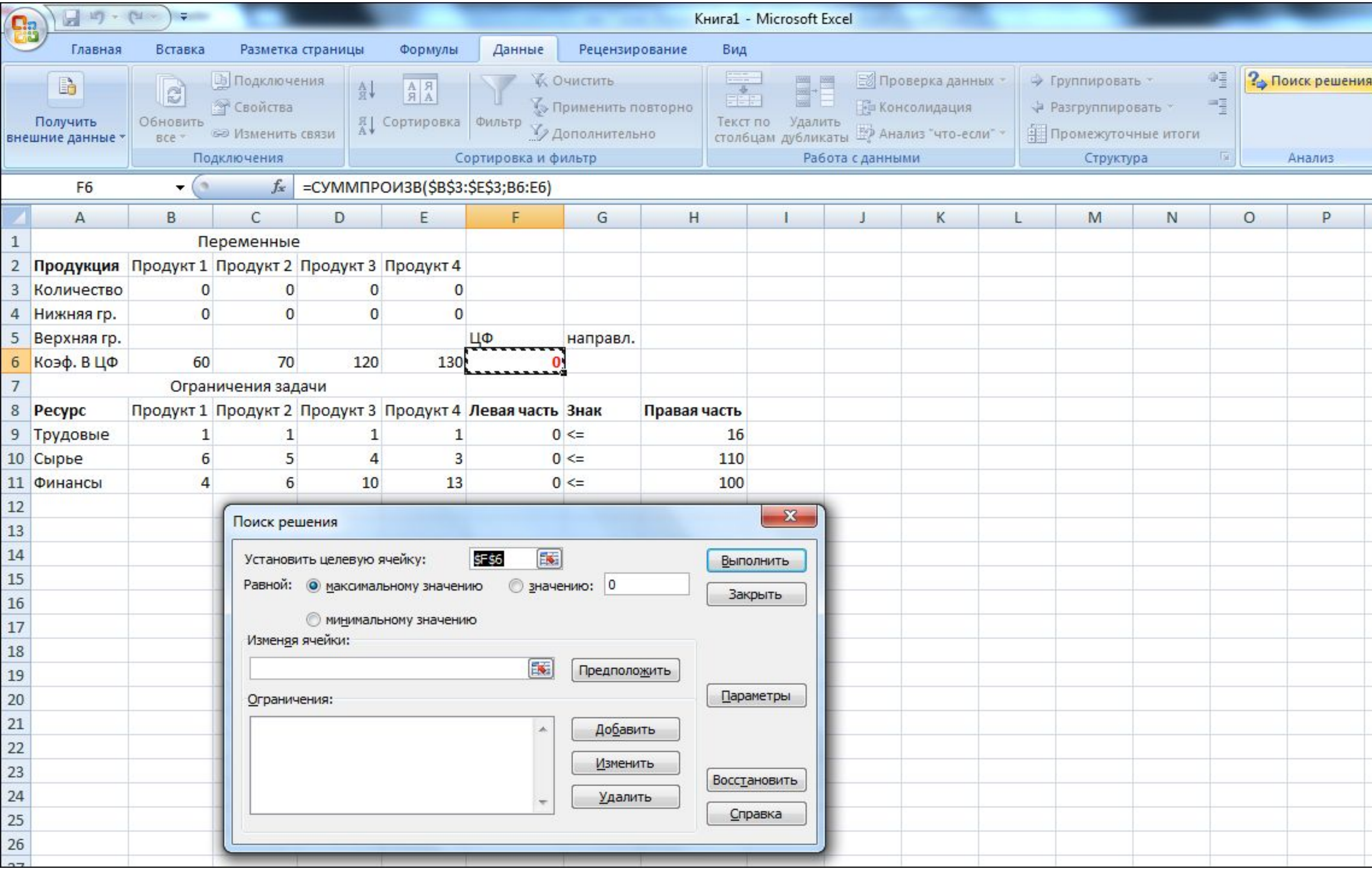
В MS Excel 2007 надстройка добавляется следующим образом.


Книга1 - Microsoft Excel
Главная
Вставка
Разметка страницы
Формулы
Данные
Рецензирование
Вид
Calibri
Вставить
Буфер обмена
A1
Параметры Excel
Основные
Формулы
Правописание
Сохранение
Дополнительно
Настройка
Надстройки
Центр управления безопасностью
Ресурсы
Управление надстройками Microsoft Office.
Надстройки
Имя
Расположение
Тип
Активные надстройки приложений
Отсутствуют активные надстройки приложений
Неактивные надстройки приложений
VBA для помощника по Интернету C:\...ry\HTML.XLAM Надстройка Excel
Имя (получатели сообщений Outlook) C:\...g\FNAME.DLL Смарт-тег
Инструменты для евро C:\...OTOOL.XLAM Надстройка Excel
Колонтитулы C:\...\OFFRHD.DLL Инспектор докуме
Мастер подстановок C:\...OOKUP.XLAM Надстройка Excel
Мастер суммирования C:\...y\SUMIF.XLAM Надстройка Excel
Настраиваемые XML-данные C:\...\OFFRHD.DLL Инспектор докуме
Невидимое содержимое C:\...\OFFRHD.DLL Инспектор докуме
Пакет анализа C:\...ANALYS32.XLL Надстройка Excel
Пакет анализа - VBA C:\...PVBAEN.XLAM Надстройка Excel
Поиск решения C:\...SOLVER.XLAM Надстройка Excel
Скрытые листы C:\...\OFFRHD.DLL Инспектор докуме
Скрытые строки и столбцы C:\...\OFFRHD.DLL Инспектор докуме
Надстройки, связанные с документами
Отсутствуют надстройки, связанные с документами
Отключенные надстройки приложений
Надстройка: Поиск решения
Издатель:
Расположение: C:\Program Files\Microsoft Office\Office12\Library\SOLVER\SOLVER.XLAM
Описание: Инструмент для поиска решения уравнений и задач оптимизации
Управление: Надстройки Excel
Перейти...
OK



Главная
Вставка
Разметка страницы
Формулы
Данные
Рецензировани
Calibri
11
Общий
Вставить
Буфер обмена
A1
Надстройки
Доступные надстройки:
VBA для помощника по Интернету
Инструменты для евро
Мастер подстановок
Мастер суммирования
Пакет анализа
Пакет анализа - VBA
Поиск решения
OK
Отмена
Обзор...
Автоматизация...
Поиск решения
Инструмент для поиска решения уравнений и задач оптимизации


Книга1 - Microsoft Excel
Главная
Вставка
Разметка страницы
Формулы
Данные
Рецензирование
Вид
Получить внешние данные
Обновить все
Подключения
Свойства
Изменить связи
Подключения
Сортировка
Фильтр
Очистить
Применить повторно
Дополнительно
Сортировка и фильтр
Текст по столбцам
Удалить дубликаты
Проверка данных
Консолидация
Анализ "что-если"
Работа с данными
Группировать
Разгруппировать
Промежуточные итоги
Структура
Поиск решения
Анализ
A1


Книга1 - Microsoft Excel
Главная
Вставка
Разметка страницы
Формулы
Данные
Рецензирование
Вид
Получить внешние данные
Обновить все
Подключения
Свойства
Изменить связи
Подключения
Сортировка
Фильтр
Очистить
Применить повторно
Дополнительно
Сортировка и фильтр
Текст по столбцам
Удалить дубликаты
Проверка данных
Консолидация
Анализ "что-если"
Работа с данными
Группировать
Разгруппировать
Промежуточные итоги
Структура
Поиск решения
Анализ
F6
=СУММПРОИЗВ($B$3:$E$3;B6:E6)
Переменные
Продукция
Продукт 1
Продукт 2
Продукт 3
Продукт 4
Количество 0 0 0 0
Нижняя гр. 0 0 0 0
Верхняя гр.
ЦФ
направл.
Коэф. В ЦФ 60 70 120 130 0
Ограничения задачи
Ресурс
Продукт 1
Продукт 2
Продукт 3
Продукт 4
Левая часть
Знак
Правая часть
Трудовые 1 1 1 1 0 <= 16
Сырье 6 5 4 3 0 <= 110
Финансы 4 6 10 13 0 <= 100
Поиск решения
Установить целевую ячейку:
$F$6
Равной:
максимальному значению
значению:
0
минимальному значению
Изменяя ячейки:
Предположить
Ограничения:
Добавить
Изменить
Удалить
Выполнить
Закрыть
Параметры
Восстановить
Справка

Поясним смысл элементов окна **Поиск решения**.

Установить целевую ячейку - определяет целевую ячейку, значение которой необходимо максимизировать или минимизировать, или сделать равным конкретному значению.

Изменяя ячейки - определяет изменяемые ячейки (искомые). Изменяемая ячейка - это ячейка, которая может быть изменена в процессе **Поиска решения** для достижения нужного результата в ячейке из окна **Установить целевую ячейку** с удовлетворением поставленных ограничений.

Предположить - отыскивает все неформульные ячейки, прямо или непрямо зависящие от формулы в окне **Установить целевую ячейку**, и помещает их ссылки в окно **Изменяя ячейки**.

Ограничения - перечисляет текущие ограничения в данной проблеме.

Добавить - выводит окно диалога “Добавить ограничение”, в котором можно добавить ограничения к текущей проблеме.

Изменить - выводит окно диалога “Изменить ограничение”, в котором можно модифицировать имеющиеся ограничения.

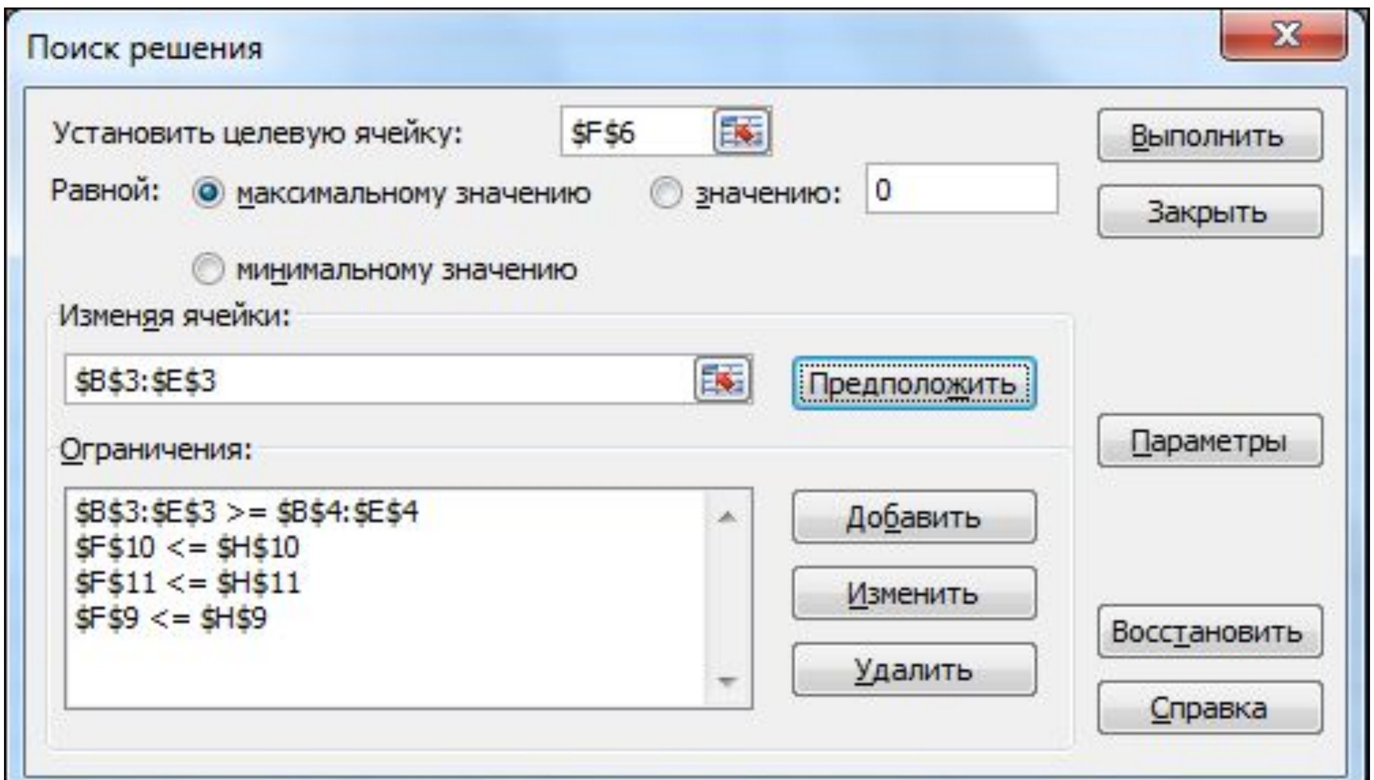
Удалить - удалить выделенное ограничение.

Выполнить - запускает процесс решения определенной проблемы.

Закрыть - закрывает окно диалога.

Параметры - выводит окно диалога “Параметры поиска решения”.

Восстановить - очищает все текущие установки проблемы и возвращает все параметры к их значениям по умолчанию.



	A	B	C	D	E	F	G	H	I
1		Переменные							
2	**Продукция**	Продукт 1	Продукт 2	Продукт 3	Продукт 4				
3	Количество	10	0	6	0				
4	Нижняя гр.	0	0	0	0				
5	Верхняя гр.					ЦФ	направл.		
6	Коэф. В ЦФ	60	70	120	130	1320			
7		Ограничения задачи							
8	**Ресурс**	Продукт 1	Продукт 2	Продукт 3	Продукт 4	**Левая часть**	**Знак**	**Правая часть**	
9	Трудовые	1	1	1	1	16	<=	16	
10	Сырье	6	5	4	3	84	<=	110	
11	Финансы	4	6	10	13	100	<=	100	

Результаты поиска решения

Решение найдено. Все ограничения и условия оптимальности выполнены.

Тип отчета

Результаты
Устойчивость
Пределы

◉ Сохранить найденное решение
○ Восстановить исходные значения

OK | Отмена | Сохранить сценарий... | Справка