CAPÍTULO 7
EJEMPLOS COMPLEMENTARIOS

1. INTRODUCCIÓN

A continuación, para una mejor comprensión de la teoría explicada en algunos pasajes anteriores de nuestro libro y su aplicación, se desarrollan en el presente capítulo algunos ejemplos complementarios adaptados respectivamente (ejemplos 2, 3 y 4), con las variaciones oportunas, de las excelentes obras de ARENALES, JOHNSTON y SPIEGEL, citadas en la bibliografía. Se trata de parcelas o situaciones geomorfológicamente atípicas pero de posible presentación en la práctica cotidiana de la ingeniería del movimiento de tierras.

2. EJEMPLO 1

Sea una porción de terreno destinada a un tramo determinado de zanja, de 1’30 m. de anchura y unos 20 metros de longitud, para la colocación de una tubería de impulsión de agua de riego de fundición dúctil y ∅ 400 mm. interior para el llenado de un lago artificial, con una pendiente variable que se pretende uniformizar, en la cual se han tomado las coordenadas relativas de los 10 puntos que se expresan en la tabla siguiente:

Tabla 1. Coordenadas de los vértices de la parcela-zanja.

VÉRTICES	COORDENADAS RELATIVAS		
	$X = X_1$ (m.)	$Y = X_2$ (m.)	$Z = Y$ (m.)
1	35,00	6,20	5,90
2	36,30	6,70	6,40
3	37,50	6,60	5,50
4	40,25	6,20	6,40
5	42,30	7,20	6,10
6	44,30	6,25	5,50
7	46,30	6,00	5,10
8	49,00	6,20	7,10
9	52,60	7,20	10,30
10	54,50	6,60	10,00

A continuación, puede verse la planta altimétrica correspondiente, con una equidistancia de las curvas de nivel de 0’50 m.:

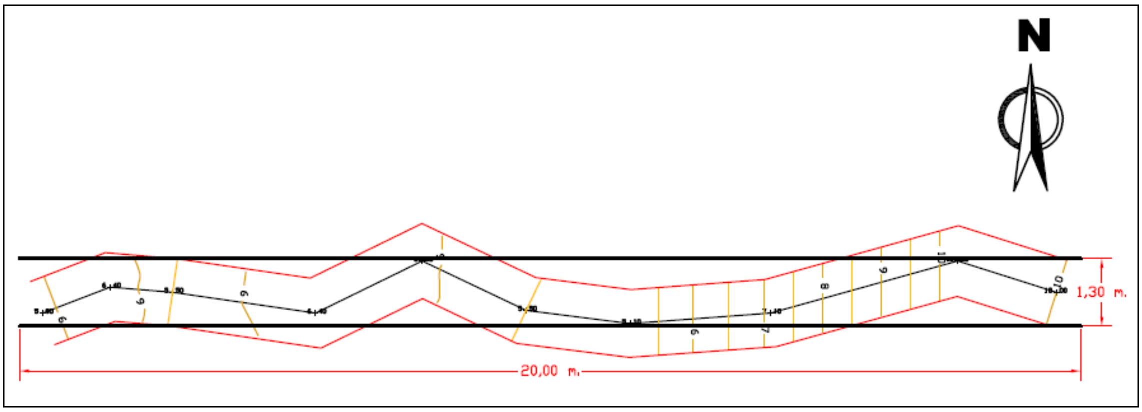

Fig. 1. Planta curvada de la zanja.

También puede verse el perfil longitudinal por el eje determinado (línea quebrada) por los puntos levantados de la zanja, con la rasante definitiva correspondiente:

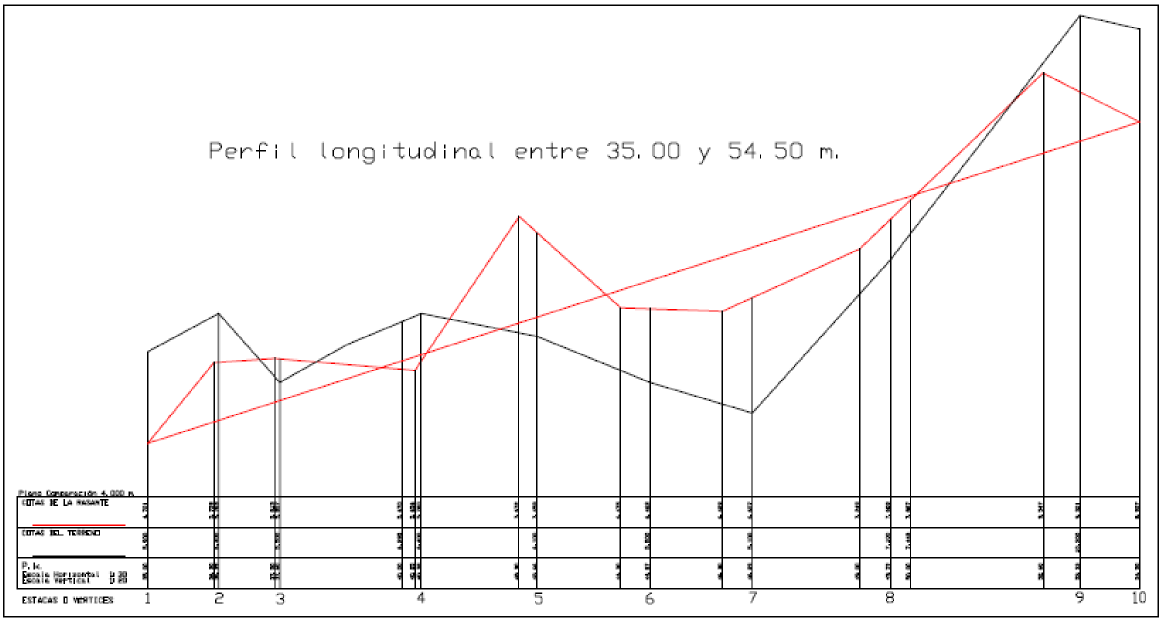

Fig. 2. Perfil longitudinal de la zanja.

a) Estímese la ecuación de regresión lineal múltiple que determina las cotas del plano óptimo definitivo de ajuste por el método tradicional de los mínimos cuadrados.

SOLUCIÓN.

Puesto que n = 10 vértices o estacas, formaremos la siguiente tabla auxiliar de cálculo:

Tabla 2. Tabla auxiliar de cálculo (I).

Estaca	Y	X_1	X_2	Y^2	X_1^2	X_2^2	$Y \cdot X_1$	$Y \cdot X_2$	$X_1 \cdot X_2$
1	5,90	35,00	6,20	34,81	1.225,00	38,44	206,50	36,58	217,00
2	6,40	36,30	6,70	40,96	1.327,69	44,89	232,32	42,88	243,21
3	5,50	37,50	6,60	30,25	1.406,25	43,56	206,25	36,30	247,50
4	6,40	40,25	6,20	40,96	1.620,06	38,44	257,60	39,68	249,55
5	6,10	42,30	7,20	37,21	1.789,29	51,84	258,03	43,92	304,56
6	5,50	44,30	6,25	30,25	1.962,49	39,06	243,65	34,38	276,88
7	5,10	46,30	6,00	26,01	2.143,69	36,00	236,13	30,60	277,80
8	7,10	49,00	6,20	50,41	2.401,00	38,44	347,90	44,02	303,80
9	10,30	52,60	7,20	106,09	2.766,76	51,84	541,78	74,16	378,72
10	10,00	54,50	6,60	100,00	2.970,25	43,56	545,00	66,00	359,70
TOTAL	**68,30**	**438,05**	**65,15**	**496,95**	**19.602,48**	**426,07**	**3.075,16**	**448,52**	**2.858,72**

Las ecuaciones normales serán:

$$\begin{cases} \Sigma Y = b_{1\cdot 23} \cdot n + b_{12\cdot 3} \cdot \Sigma X_1 + b_{13\cdot 2} \cdot \Sigma X_2 \\ \Sigma Y \cdot X_1 = b_{1\cdot 23} \cdot \Sigma X_1 + b_{12\cdot 3} \cdot \Sigma X_1^2 + b_{13\cdot 2} \cdot \Sigma X_1 \cdot X_2 \\ \Sigma Y \cdot X_2 = b_{1\cdot 23} \cdot \Sigma X_2 + b_{12\cdot 3} \cdot \Sigma X_1 \cdot X_2 + b_{13\cdot 2} \cdot \Sigma X_2^2 \end{cases}$$

que, substituyendo los valores obtenidos en la tabla anterior resulta:

$$\begin{cases} 68'30 = b_{1\cdot 23} \cdot 10 + b_{12\cdot 3} \cdot 438'05 + b_{13\cdot 2} \cdot 65'15 \\ 3.075'16 = b_{1\cdot 23} \cdot 438'05 + b_{12\cdot 3} \cdot 19.602'48 + b_{13\cdot 2} \cdot 2.858'72 \\ 448'52 = b_{1\cdot 23} \cdot 65'15 + b_{12\cdot 3} \cdot 2.858'72 + b_{13\cdot 2} \cdot 426'07 \end{cases}$$

Se trata de un sistema no homogéneo (heterogéneo), de 3 ecuaciones con 3 incógnitas, $|A| \neq 0$, compatible y determinado, con solución única que podemos hallar mediante la aplicación del método de la inversión de la matriz, el de la triangularización de Gauss-Jordan o bien la regla de Cramer (puesto que la matriz de los coeficientes de las incógnitas es invertible). Aplicando este último procedimiento, se tiene:

$$b_{1\cdot 23} = \frac{\begin{vmatrix} 68'30 & 438'05 & 65'15 \\ 3.075'16 & 19.602'48 & 2.858'72 \\ 448'52 & 2.858'72 & 426'07 \end{vmatrix}}{\begin{vmatrix} 10 & 438'05 & 65'15 \\ 438'05 & 19.602'48 & 2.858'72 \\ 65'15 & 2.858'72 & 426'07 \end{vmatrix}} = \frac{-76.787'81}{6.459'886} = -11'887\,;$$

$$b_{12\cdot 3} = \frac{\begin{vmatrix} 10 & 68'30 & 65'15 \\ 438'05 & 3.075'16 & 2.858'72 \\ 65'15 & 448'52 & 426'07 \end{vmatrix}}{|A|} = \frac{1.176'186}{6.459'886} = 0'182\,;$$

$$b_{13\cdot 2} = \frac{\begin{vmatrix} 10 & 438'05 & 68'30 \\ 438'05 & 19.602'48 & 3.075'16 \\ 65'15 & 2.858'72 & 448'52 \end{vmatrix}}{|A|} = \frac{10.650'158}{6.459'886} = 1'649\,;$$

De este modo, el plano de regresión mínimocuadrática tendrá de ecuación general:

$$\mathbf{Y = -11'887 + 0'182 \cdot X_1 + 1'649 \cdot X_2}$$

b) Estímese la misma ecuación de regresión lineal múltiple del plano óptimo de nivelación por aplicación del método presentado en este libro.

SOLUCIÓN.

Si ahora, a efectos puramente comparativos, aplicamos el método de cálculo establecido por nosotros a partir de la expresada hoja *Excel* con *Solver*, se tendrá la ecuación de las cotas definitivas ($Z = T_i$):

$$\mathbf{Y = -11'856 + 0'182 \cdot X_1 + 1'643 \cdot X_2}$$

que resulta algo más ajustada que la anteriormente obtenida por el método analítico tradicional y ofrece la siguiente tabla de discrepancias o correcciones:

Tabla 3. Cotas definitivas y correcciones (I).

Estaca	**X_1**	**X_2**	**Y**	**T_i**	**$d_i = Y_i - T_i$**
1	35,00	6,20	5,9	4,701	+1,199
2	36,30	6,70	6,4	5,759	+0,641
3	37,50	6,60	5,5	5,813	-0,313
4	40,25	6,20	6,4	5,656	+0,744
5	42,30	7,20	6,1	7,672	-1,572
6	44,30	6,25	5,5	6,475	-0,975
7	46,30	6,00	5,1	6,429	-1,329
8	49,00	6,20	7,1	7,249	-0,149
9	52,60	7,20	10,3	9,547	+0,753
10	54,50	6,60	10,0	8,907	+1,093
TOTAL	**438,05**	**65,15**	**68,3**	**68,300**	**±0**

c) Calcúlese el correspondiente grado de explanación.

SOLUCIÓN.

Para tener una medida del grado de explanación, en base a lo explicitado en el capítulo 5 anterior, igualaremos a +10,00 m. la cota relativa media o centro de gravedad de la parcela en estudio, cuyo valor resulta de dividir la suma de las cotas iniciales del terreno natural por el número de vértices (68,3/10 = 6'83 m.), con lo que se tendrá la siguiente tabla:

Vértices	**Cotas relativas iniciales (Y_i)**	**Cotas relativas definitivas (T_i)**	**d_i (Y_i - T_i)**	**d_i^2**	**d_i^2/T_i**
1	9,07	7,871	1,199	1,439	0,183
2	9,57	8,929	0,641	0,411	0,046
3	8,67	8,983	-0,313	0,098	0,011
4	9,57	8,826	0,744	0,553	0,063
5	9,27	10,842	-1,572	2,472	0,228
6	8,67	9,645	-0,975	0,951	0,099
7	8,27	9,599	-1,329	1,765	0,184
8	10,27	10,419	-0,149	0,022	0,002
9	13,47	12,717	0,753	0,567	0,045
10	13,17	12,077	1,093	1,195	0,099
Σ	100,00	100,000	$\pm$0,000	9,474	**χ^2=0,959**

Obsérvese que en la tabla anterior hemos definido el $d_i = Y_i - T_i$ como diferencia entre las cotas del terreno natural o iniciales y las definitivas que se deducen de la aplicación de nuestro modelo de explanación. Ello es así con el objetivo de adecuarnos a la terminología utilizada para el cálculo de chi-cuadrado que realizaremos a continuación.

El error estándar o típico de la estima de esta regresión múltiple (triple) vendrá dado por la expresión (véase el capítulo 5) con m = 2 variables explicativas correspondientes a la abscisa y la ordenada de cada punto:

$$S_{xy} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - T_i)^2}{N - m - 1}} = \sqrt{\frac{9{,}474}{10 - 2 - 1}} = 1'16\,\text{m.}$$

Para N – 1 = 10 – 1 = 9 grados de libertad, se tiene un $\chi^2_{0,5}$ = 1'735, buscando en la tabla de percentiles de la distribución teórica de probabilidad chi-cuadrado que figura en el anexo 3. Al ser: χ^2=0'959 < 1'735 puede considerarse aceptable el volumen de explanación a realizar en la parcela que nos ocupa en base al estadígrafo utilizado.

De haberse considerado alternativamente: N – m – 1 = 7 g.l., se obtendría un valor más exigente de $\chi^2_{0,5}$ = 0’989 > 0’959, lo que siquiera por bien poco no modificaría tampoco las conclusiones anteriormente obtenidas.

Por otra parte, el “grado de explanación” determinado, como ya se ha visto, por el “coeficiente de contingencia” C derivado de la distribución de probabilidad chi-cuadrado (χ^2), vendrá dado por la expresión:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{0'959}{0'959 + 10}} = 0'30 \cong 30\%,$$

siendo N = 10 el número de estacas o vértices de nivelación de la zanja considerado.

3. EJEMPLO 2

Sea ahora una parcela o porción de terreno en talud o ladera de montaña de gran desnivel, en la cual se han tomado las coordenadas relativas de 10 puntos que se expresan en la tabla siguiente y que, para mayor simplificación del cálculo, se han obviado las cifras decimales ajustándolas a las enteras:

Tabla 4. Coordenadas de los vértices de la parcela (II).

VÉRTICES	**COORDENADAS RELATIVAS**		
	X = X_1 (m.)	**Y = X_2 (m.)**	**Z = Y (m.)**
1	86	96	27
2	29	71	7
3	61	120	24
4	67	77	20
5	45	80	8
6	9	89	5
7	43	107	17
8	37	81	9
9	17	65	3
10	71	102	35

A continuación, puede verse el plano altimétrico curvado correspondiente, con algunos perfiles (1 longitudinal y 3 transversales) del mismo y una equidistancia vertical entre las curvas de nivel de 1’00 m.:

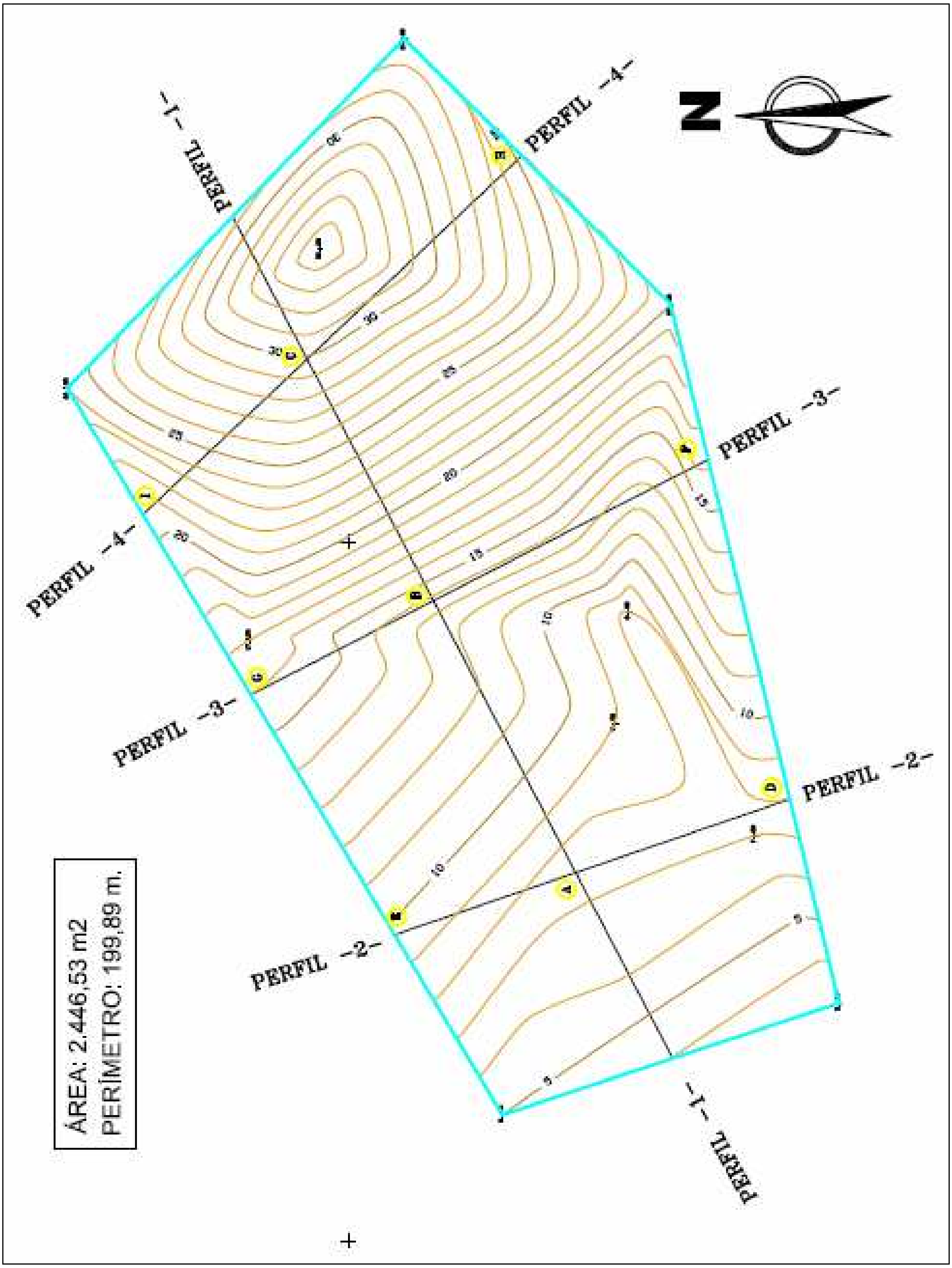

Fig. 3. Planta curvada de la parcela (I).

a) Estimación de las medias aritméticas, varianzas, covarianzas y coeficientes de correlación lineal simples.

SOLUCIÓN.

De estos datos se deducen las siguientes características de la distribución conjunta de las tres variables topográficas analizadas: X_1, X_2 e Y:

-Medias aritméticas de las variables:

$$\begin{cases} \overline{X}_1 = \dfrac{\sum X_1}{n} = \dfrac{465}{10} = 46'5 \text{ m.} \\ \overline{X}_2 = \dfrac{\sum X_2}{n} = \dfrac{888}{10} = 88'8 \text{ m.} \\ \overline{Y} = \dfrac{\sum Y}{n} = \dfrac{155}{10} = 15'5 \text{ m.} \end{cases}$$

-Varianzas y desviaciones típicas o "standard" de las variables:

$$\begin{cases} s_1^2 = \dfrac{\sum X_1^2}{n} - \overline{X}_1^2 = \dfrac{27.101}{10} - 46'5^2 = 547'85 \text{ m}^2; \quad s_1 = 23'41 \text{ m.} \\ s_2^2 = \dfrac{\sum X_2^2}{n} - \overline{X}_2^2 = \dfrac{81.546}{10} - 88'8^2 = 269'16 \text{ m}^2; \quad s_2 = 16'41 \text{ m.} \\ s_y^2 = \dfrac{\sum Y^2}{n} - \overline{Y}^2 = \dfrac{3.447}{10} - 15'5^2 = 104'45 \text{ m}^2; \quad s_y = 10'22 \text{ m.} \end{cases}$$

-Covarianzas de las variables:

$$\begin{cases} s_{12} = \dfrac{\sum X_1 X_2}{n} - \overline{X}_1 \cdot \overline{X}_2 = \dfrac{43.140}{10} - 46'5 \times 88'8 = 184'80 \text{ m}^2 \\ s_{1y} = \dfrac{\sum X_1 Y}{n} - \overline{X}_1 \cdot \overline{Y} = \dfrac{9.334}{10} - 46'5 \times 15'5 = 212'65 \text{ m}^2 \\ s_{2y} = \dfrac{\sum X_2 Y}{n} - \overline{X}_2 \cdot \overline{Y} = \dfrac{14.907}{10} - 88'8 \times 15'5 = 114'30 \text{ m}^2 \end{cases}$$

-Coeficientes de correlación lineal simple:

$$\begin{cases} r_{12} = \dfrac{s_{12}}{s_1 \cdot s_2} = \dfrac{184'80}{23'41 \times 16'41} = 0'4811 \\[2ex] r_{1y} = \dfrac{s_{1y}}{s_1 \cdot s_y} = \dfrac{212'65}{23'41 \times 10'22} = 0'8888 \\[2ex] r_{2y} = \dfrac{s_{2y}}{s_2 \cdot s_y} = \dfrac{114'30}{16'41 \times 10'22} = 0'6815 \end{cases}$$

Como el coeficiente de correlación lineal simple que relaciona las variables X_1 y X_2 no es muy alto (r_{12} = 0’4811), no existe multicolinealidad[1] y ello favorece la estimación de la ecuación de regresión de la cota Y sobre las coordenadas X_1 y X_2.

Para la realización de las anteriores determinaciones se habrá elaborado, previamente, el cuadro siguiente:

Tabla 5. Tabla auxiliar de cálculo (II).

Estaca	X_1	X_2	Y	$X_1 \cdot X_2$	$X_1 \cdot Y$	$X_2 \cdot Y$	X_1^2	X_2^2	Y^2
1	86	96	27	8.256	2.322	2.592	7.396	9.216	729
2	29	71	7	2.059	203	497	841	5.041	49
3	61	120	24	7.320	1.464	2.880	3.721	14.400	576
4	67	77	20	5.159	1.340	1.540	4.489	5.929	400
5	45	80	8	3.600	360	640	2.025	6.400	64
6	9	89	5	801	45	445	81	7.921	25
7	43	107	17	4.601	731	1.819	1.849	11.449	289
8	37	81	9	2.997	333	729	1.369	6.561	81
9	17	65	3	1.105	51	195	289	4.225	9
10	71	102	35	7.242	2.485	3.570	5.041	10.404	1.225
TOTAL	**465**	**888**	**155**	**43.140**	**9.334**	**14.907**	**27.101**	**81.546**	**3.447**

[1] El proceso o término antedicho de “multicolinealidad” es una situación en la que se presenta una fuerte correlación entre las variables explicativas del modelo, o sea, las coordenadas de los puntos analizados. La correlación ha de ser fuerte, ya que siempre existirá correlación entre dos variables explicativas en un modelo, es decir, la no correlación de dos variables es un proceso idílico, que sólo se podría encontrar en condiciones de laboratorio. Originalmente, el término de multicolinealidad significó la existencia de una relación perfecta o exacta entre las variables explicativas de un modelo de regresión. En la actualidad se incluye en la multicolinealidad el término de “error estocástico”. La multicolinealidad así expuesta se refiere solamente a relaciones lineales entre las variables, pero no elimina las relaciones no lineales existentes entre ellas. Se supone que en un modelo clásico de regresión lineal no hay multicolinealidad debido a que si la multicolinealidad es perfecta los coeficientes de la regresión de las variables son indeterminados y sus errores estándar son infinitos. Si la multicolinealidad es menos que perfecta los coeficientes de regresión poseen grandes errores estándar, lo que provoca que los coeficientes no puedan ser estimados con gran precisión.

b) Estímese la ecuación de regresión lineal múltiple que determina las cotas del plano óptimo definitivo de ajuste. Calcúlense las correcciones a efectuar en los diversos vértices o estacas del terreno para establecer las cotas definitivas del mismo, así como el "grado de explanación" resultante.

SOLUCIÓN.

La ecuación de regresión lineal múltiple que tratamos de estimar posee, como sabemos, la configuración analítica siguiente:

$$Y = a + b_1 X_1 + b_2 X_2$$

El *coeficiente de regresión parcial* de Y sobre X_1 puede estimarse con la fórmula:

$$\hat{b}_1 = \frac{s_y}{s_1} \cdot \frac{r_{1y} - r_{2y} r_{12}}{1 - r_{12}^2} = \frac{10'22}{23'41} \cdot \frac{0'8888 - 0'6815 \times 0'4811}{1 - 0'4811^2} =$$

$$= \frac{10'22}{23'41} \cdot \frac{0'5609}{0'7685} = 0'3186$$

De la misma manera, el coeficiente de regresión de Y sobre X_2 se define por:

$$\hat{b}_2 = \frac{s_y}{s_2} \cdot \frac{r_{2y} - r_{1y} r_{12}}{1 - r_{12}^2} = \frac{10'22}{16'41} \cdot \frac{0'6815 - 0'8888 \times 0'4811}{1 - 0'4811^2} =$$

$$= \frac{10'22}{16'41} \cdot \frac{0'2539}{0'7685} = 0'2058$$

El parámetro independiente *a* se deduce de la igualdad:

$$Y - \overline{Y} = \hat{b}_1 (X_1 - \overline{X}_1) + \hat{b}_2 (X_2 - \overline{X}_2) \ ,$$

de donde,

$$Y - 15'5 = 0'3186\ (X_1 - 46'5) + 0'2058\ (X_2 - 88'8) \ ,$$

$$Y - 15'5 = 0'3186\ X_1 - 14'8149 + 0'2058\ X_2 - 18'2750$$

y resulta: $\mathbf{Y = -17'5899 + 0'3186\ X_1 + 0'2058\ X_2}$

que es la ecuación de regresión pedida y en la que el término independiente toma el valor,

$$a = -17'5899 \cong -17'59$$

Para calcular las discrepancias existentes entre los valores observados de las 10 cotas del terreno y los teóricos deducidos de la ecuación de regresión lineal múltiple para conformar el plano óptimo de nivelación, deben calcularse "grosso modo" las diferencias:

$$d_i = Y_i - T_i = Y_i - (-17'59 + 0'32\ X_1 + 0'21\ X_2)$$

lo que se consigue dando a cada terna de valores de las coordenadas (X_1, X_2, Y) los que corresponden a cada una de las diez estacas seleccionadas. Puede comprobarse (tabla 6) que:

$$\begin{cases} \Sigma Y_i = \Sigma T_i = 155 \\ \Sigma d_i = 0\ , \end{cases}$$

como se sabe que teóricamente debe ocurrir.

Como consecuencia de la realización de los cálculos correspondientes, se obtendrá la siguiente tabla:

Tabla 6. Cotas definitivas y correcciones (II).

Estaca	**X_1**	**X_2**	**Y**	**T_i**	**$d_i = Y_i - T_i$**
1	86	96	27	30	-3
2	29	71	7	6	+1
3	61	120	24	26	-2
4	67	77	20	20	0
5	45	80	8	13	-5
6	9	89	5	4	+1
7	43	107	17	18	-1
8	37	81	9	11	-2
9	17	65	3	1	+2
10	71	102	35	26	+9
TOTAL	**465**	**888**	**155**	**155**	**±0**

Obsérvese que, en este caso, hemos calculado el plano óptimo de regresión siguiendo pautadamente el procedimiento teórico clásico, aunque resulta mucho más operativo hacerlo mediante el modelo que presenta la hoja de cálculo *Excel* con *Solver* que presentamos en otros apartados de este mismo libro. Así pues, si ahora, a efectos puramente comparativos, aplicamos el método de cálculo establecido por nosotros a partir de la expresada hoja *Excel*, se tendrá la ecuación de las cotas definitivas ($Z = T_i$):

$$\mathbf{Y = -17'5979 + 0'3187\ X_1 + 0'2058\ X_2}$$

que ofrece la siguiente tabla, con alguna mayor precisión que la anterior:

Tabla 7. Cotas definitivas y correcciones (III).

Estaca	**X_1**	**X_2**	**Y**	**T_i**	**$d_i = Y_i - T_i$**
1	86	96	27	29,567	-2,567
2	29	71	7	6,256	+0,744
3	61	120	24	26,539	-2,539
4	67	77	20	19,602	+0,398
5	45	80	8	13,208	-5,208
6	9	89	5	3,587	+1,413
7	43	107	17	18,127	-1,127
8	37	81	9	10,864	-1,864
9	17	65	3	1,197	+1,803
10	71	102	35	26,021	+8,979
TOTAL	**465**	**888**	**155**	**155,000**	**±0**

Para tener una medida del grado de explanación, en base a lo explicitado en el capítulo 5 anterior, igualaremos a +10,00 m. la cota relativa media o centro de gravedad de la parcela en estudio, cuyo valor resulta de dividir la suma de las cotas iniciales del terreno natural por el número de vértices (155/10 = 15’50 m.). En este caso se presentarán cotas relativas definitivas negativas, con lo que para evitar distorsiones en el proceso de cálculo deberán considerarse los cocientes d_i^2/T_i de la última columna de la tabla siguiente en valor absoluto, con lo que se tendrá:

Vértices	**Cotas relativas iniciales (Y_i)**	**Cotas relativas definitivas (T_i)**	**d_i ($Y_i - T_i$)**	**d_i^2**	**$\lvert d_i^2/T_i \rvert$**
1	21,50	24,067	-2,567	6,590	0,274
2	1,50	0,756	0,744	0,553	0,732
3	18,50	21,039	-2,539	6,446	0,306
4	14,50	14,102	0,398	0,159	0,011
5	2,50	7,708	-5,208	27,119	3,518
6	-0,50	-1,913	1,413	1,998	1,044
7	11,50	12,627	-1,127	1,270	0,101
8	3,50	5,364	-1,864	3,474	0,648
9	-2,50	-4,303	1,803	3,251	0,755
10	29,50	20,521	8,979	80,615	3,928
Σ	100,00	100,000	±0,000	131,474	**χ^2=11,317**

Obsérvese que en la tabla anterior hemos definido el $d_i = Y_i – T_i$ como diferencia entre las cotas del terreno natural o iniciales y las definitivas que se deducen de la aplicación de nuestro modelo de explanación. Ello es así con el objetivo de adecuarnos a la terminología utilizada para el cálculo de chi-cuadrado que realizaremos a continuación.

El error estándar o típico de la estima de esta regresión múltiple (triple) vendrá dado por la expresión (véase el capítulo 5) con m = 2 variables explicativas correspondientes a la abscisa y la ordenada de cada punto:

$$S_{xy} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - T_i)^2}{N - m - 1}} = \sqrt{\frac{131{,}474}{10 - 2 - 1}} = 4'33\ \text{m.}$$

Para N – 1 = 10 – 1 = 9 grados de libertad, se tiene un $\chi^2_{0,5}$ = 1’735, buscando en la tabla de percentiles de la distribución teórica de probabilidad chi-cuadrado que figura en el anexo 3. Al ser: χ^2=11’317 > 1’735 puede considerarse desde luego inaceptable el volumen de explanación a realizar en la parcela que nos ocupa en base al estadígrafo utilizado.

Por otra parte, el “grado de explanación” determinado, como ya se ha visto, por el “coeficiente de contingencia” C derivado de la distribución de probabilidad chi-cuadrado (χ^2), vendrá dado por la expresión:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{11'317}{11'317 + 10}} = 0'73 \cong 73\%\,,$$

siendo N = 10 el número de estacas o vértices de nivelación considerado.

Comparando los valores obtenidos de χ^2, S_{xy} y C para esta parcela y la del ejemplo anterior, se observa que en este segundo caso la explanación a efectuar resultará mucho mayor y por tanto muy desfavorable, exigiendo un mayor movimiento de tierras, tanto de desmonte como de terraplén, con el incremento de coste correspondiente.

Obsérvese, como ya se ha señalado, que en este último caso nos habían aparecido valores de T_i negativos que hacían que la expresión de la última columna de la tabla anterior d_i^2/T_i resultase también negativa en alguna ocasión. Ello podrá suceder, aunque no siempre, en aquellos terrenos en que las diferencias de cota entre sus puntos sea considerable (normalmente superiores a 20 m.). En estos casos (como en los dos ejemplos que siguen en este mismo capítulo de nuestro libro) y como alternativa a la solución empleada de considerar los valores absolutos de los cocientes d_i^2/T_i aconsejamos, para evitar confusiones de interpretación en cuanto a la verdadera cuantía de la χ^2, considerar, por ejemplo, una cota relativa media de +20’00 m. en vez de la de +10’00 m. que hemos venido empleando hasta ahora, pero siempre en el bien entendido de que las comparaciones a efectuar entre diferentes terrenos a explanar deberán llevarse a cabo siempre sobre las mismas bases de partida (la misma cota relativa media, centroide o centro de gravedad).

c) Calcúlense los tres posibles coeficientes de correlación parcial que pueden deducirse con los resultados obtenidos e interprétense los resultados conseguidos.

SOLUCIÓN.
Los coeficientes de correlación parcial pedidos son los siguientes:

$$\left\{\begin{array}{l} r_{1y\cdot 2} = \dfrac{r_{1y} - r_{12}\cdot r_{2y}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{2y}^2}} = \dfrac{0'8888 - 0'4811 \times 0'6815}{\sqrt{1-0'4811^2}\sqrt{1-0'6815^2}} = \\ \qquad = \dfrac{0'5609}{0'8767 \times 0'7318} = 0'8742 \\ \\ r_{2y\cdot 1} = \dfrac{r_{2y} - r_{12}\cdot r_{1y}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{1y}^2}} = \dfrac{0'6815 - 0'4811 \times 0'8888}{\sqrt{1-0'4811^2}\sqrt{1-0'8888^2}} = \\ \qquad = \dfrac{0'2533}{0'8767 \times 0'4583} = 0'6304 \\ \\ r_{12\cdot y} = \dfrac{r_{12} - r_{1y}\cdot r_{2y}}{\sqrt{1-r_{1y}^2}\sqrt{1-r_{2y}^2}} = \dfrac{0'4811 - 0'8888 \times 0'6815}{\sqrt{1-0'8888^2}\sqrt{1-0'6815^2}} = \\ \qquad = \dfrac{-0'1246}{0'4583 \times 0'7318} = -0'3715 \end{array}\right.$$

Los tres coeficientes de correlación parcial de tal suerte calculados han resultado inferiores a los coeficientes de correlación lineal simple, es decir,

$$\left\{\begin{array}{l} -0'3715 < 0'4811 \\ 0'8742 < 0'8888 \\ 0'6304 < 0'6815 \end{array}\right.$$

pero en tanto que los dos últimos no difieren notablemente, la correlación entre la X_1 y la X_2 es negativa, bajo el supuesto de que se consideran constantes las cotas del terreno (Y). Esto quiere decir que para puntos del terreno con cotas taquimétricas de la misma cuantía, a mayor abscisa X_1 le corresponde menor ordenada X_2 y recíprocamente. Esta circunstancia puede apreciarse fácilmente de la contemplación del plano altimétrico en planta del terreno que se adjunta.

d) Estímese una ecuación lineal que explique la coordenada (X_1) en función de la cota taquimétrica (Y) y de la coordenada (X_2) con los datos obtenidos, pero haciendo uso de la definición de un coeficiente de correlación parcial como una media geométrica de dos coeficientes de regresión múltiple lineal.

SOLUCIÓN.
La ecuación pedida puede expresarse en la forma:

$$X_1 = a' + b'_1\, Y + b'_2\, X_2$$

que al compararla con la ecuación de regresión lineal múltiple del plano óptimo de nivelación:

$$Y = a + b_1\, X_1 + b_2\, X_2$$

nos permite escribir:

$$r_{1y\cdot 2} = \sqrt{\hat{b}_1 \cdot \hat{b}'_1}$$

De aquí se deduce que:

$$r^2_{1y\cdot 2} = \hat{b}_1 \cdot \hat{b}'_1 \text{ ; o sea:}$$

$$\hat{b}'_1 = \frac{r^2_{1y\cdot 2}}{\hat{b}_1} = \frac{0'8742^2}{0'3186} = 2'3987$$

La estimación de $\hat{b}'_2$ debe realizarse por el método ordinario, es decir, aplicando la fórmula:

$$b'_2 = \frac{s_1}{s_2} \cdot \frac{r_{21} - r_{y1} r_{y2}}{1 - r^2_{y2}} = \frac{23'41}{16'41} \cdot \frac{0'4811 - 0'8888 \times 0'6815}{1 - 0'6815^2} =$$

$$= -\frac{23'41}{16'41} \cdot \frac{0'1246}{0'5356} = -0'3319$$

esto es,

$$X_1 - 46'5 = 2'3987\,(Y - 15'5) - 0'3319\,(X_2 - 88'8)\ ,$$

y despejando la X_1 se obtiene:

$$\mathbf{X_1 = 38'7928 + 2'3987\, Y - 0'3319\, X_2} \qquad (1)$$

que es la ecuación pedida.

Si ahora despejamos X_1 en la ecuación de regresión inicialmente obtenida en el apartado b) anterior, se tiene:

$$Y = - 17'5899 + 0'3186 \cdot X_1 + 0'2058 \cdot X_2$$

esto es:

$$3'138732 \cdot Y = - 55'21 + X_1 + 0'645951 \cdot X_2 ,$$

con lo que:

$$\mathbf{X_1 = 55'21 + 3'138732 \cdot Y - 0'645951 \cdot X_2} \qquad (2)$$

Los resultados de ambas determinaciones varían sensiblemente para cada estaca como puede comprobarse en la tabla siguiente, aunque la suma total de las diferencias se anula.

En efecto, dando valores a X_2 y a Y, se tienen unos valores de X_1 según las opciones (1) o (2) empleadas de:

Tabla 8. Tabla auxiliar de cálculo (III).

VÉRTICES	Y	X_2	X_1 (1)	X_1 (2)	$X_1(1)-X_1(2)$
1	27	96	71,6953	77,944468	-6,249168
2	7	71	32,0188	31,318603	0,700197
3	24	120	56,5336	53,025448	3,508152
4	20	77	61,2105	68,246413	-7,035913
5	8	80	31,4304	28,643776	2,786624
6	5	89	21,2472	13,414021	7,833179
7	17	107	44,0574	39,451687	4,605713
8	9	81	33,4972	31,136557	2,360643
9	3	65	24,4154	22,639381	1,776019
10	35	102	88,8935	99,178618	-10,285118
TOTAL			**464,9993**	**464,998972**	**0,000328≈0**

e) Con los datos de la tabla 6 estímese la varianza residual de la cota taquimétrica del terreno (Y) sobre la abscisa (X_1) y la ordenada (X_2). De este resultado y la varianza de Y (s^2_y = 104'45 m^2) dedúzcanse los coeficientes de determinación y correlación múltiple lineal de Y sobre las coordenadas X_1 y X_2.

SOLUCIÓN.

Como la varianza residual[2] es igual a:

$$s^2_{ry\cdot 12} = \frac{\sum d_i^2}{n}$$

y d_i toma los valores {-3, 1, -2, 0, -5, 1, -1, -2, 2, 9}, como puede comprobarse de la tabla 6, también resultará que:

Σd_i^2 = 9 + 1 + 4 + 0 + 25 + 1 + 1 + 4 + 4 + 81 = 130 m^2 , de donde:

$$s^2_{ry\cdot 12} = \frac{130}{10} = 13 \text{ m}^2$$

Por otra parte, el coeficiente de determinación o crítico pedido se calcula según la fórmula:

$$R^2_{y\cdot 12} = 1 - \frac{s^2_{ry\cdot 12}}{s^2_y} = 1 - \frac{13}{104'45} = 0'8755 \text{ ,}$$

y el de correlación múltiple lineal es la raíz cuadrada del anterior coeficiente de determinación, es decir,

$$R_{y\cdot 12} = \sqrt{0'8755} = 0'9357$$

que resulta ciertamente elevado. En efecto, el coeficiente de correlación múltiple se encuentra comprendido entre 0 y 1. Cuanto más se acerque a 1 mejor es la relación lineal existente entre las variables (coordenadas) del problema; por el contrario, cuanto más se acerque a 0 la relación lineal será tanto peor. Si su valor es 1, la correlación se denomina "perfecta".

Veamos, en fin, que aunque un coeficiente de correlación con valor 0 indica que no existe relación lineal entre las coordenadas de los puntos del terreno analizado, es posible que pueda existir entre ellas una relación no lineal, lo que obligaría a enfocar la resolución del problema desde el punto de vista del ajuste a una superficie de nivelación de segundo orden, con lo que la solución resultante ya no sería un plano de

[2] Es la que queda cuando de la varianza total se extraen las dos varianzas sistemáticas identificadas. La raíz cuadrada de la varianza residual es conocida como el *error típico de la estima*, que ya venimos aplicando en nuestro trabajo. La varianza residual coincide también con la suma de cuadrados de las diferencias entre los valores de la variable dependiente observados y estimados por la función de regresión, dividiendo el resultado final por el tamaño empleado de la muestra al que debe restársele el número de variables explicativas del modelo.

nivelación o explanación sino una superficie curva o alabeada (ver el capítulo siguiente de nuestro libro y el anexo nº: 2).

f) Calcúlese de nuevo el coeficiente de determinación obtenido en el apartado anterior en función de los coeficientes de correlación lineal simple estimados anteriormente.

SOLUCIÓN.
La fórmula más comúnmente empleada para calcular un coeficiente de determinación de la variable Y sobre las restantes coordenadas X_1 y X_2, que da información respecto al grado de interdependencia de las tres variables relacionadas, es la siguiente:

$$R^2_{y\cdot 12} = \frac{r^2_{1y} + r^2_{2y} - 2r_{1y}r_{2y}r_{12}}{1 - r^2_{12}};$$

aplicando los datos que ya teníamos y substituyendo en la anterior expresión, se obtiene:

$$R^2_{y\cdot 12} = \frac{0'8888^2 + 0'6815^2 - 2 \times 0'8888 \times 0'6815 \times 0'4811}{1 - 0'4811^2} =$$
$$= \frac{0'7900 + 0'4644 - 0'5828}{1 - 0'2315} = \frac{0'6716}{0'7685} = 0'8739$$

que difiere sólo en 16 diezmilésimas del calculado por el procedimiento anterior, porque en aquél se utilizaron unos valores poco precisos de las discrepancias d_i. Con ello, el coeficiente de correlación múltiple lineal definitivo vendrá dado por:

$$R_{y\cdot 12} = \sqrt{0'8739} = 0'9348$$

g) Obténgase la expresión matricial del modelo de regresión lineal múltiple propuesto.

SOLUCIÓN.
La ecuación considerada, como ya se ha visto, es la siguiente:

$$Y = a + b_1 \cdot X_1 + b_2 \cdot X_2 ;$$

Pues bien, al dar a Y, X_1, y X_2 los valores que figuran en la tabla 4 se obtienen las diez igualdades siguientes:

$$\begin{cases} 27 = a + 86\ b_1 + 96\ b_2 \\ 7 = a + 29\ b_1 + 71\ b_2 \\ 24 = a + 61\ b_1 + 120\ b_2 \\ 20 = a + 67\ b_1 + 77\ b_2 \\ 8 = a + 45\ b_1 + 80\ b_2 \\ 5 = a + 9\ b_1 + 89\ b_2 \\ 17 = a + 43\ b_1 + 107\ b_2 \\ 9 = a + 37\ b_1 + 81\ b_2 \\ 3 = a + 17\ b_1 + 65\ b_2 \\ 35 = a + 71\ b_1 + 102\ b_2 \end{cases}$$

Este conjunto de igualdades puede también presentarse en forma matricial del siguiente modo:

$$\begin{bmatrix} 27 \\ 7 \\ 24 \\ 20 \\ 8 \\ 5 \\ 17 \\ 9 \\ 3 \\ 35 \end{bmatrix} = \begin{bmatrix} 1 & 86 & 96 \\ 1 & 29 & 71 \\ 1 & 61 & 120 \\ 1 & 67 & 77 \\ 1 & 45 & 80 \\ 1 & 9 & 89 \\ 1 & 43 & 107 \\ 1 & 37 & 81 \\ 1 & 17 & 65 \\ 1 & 71 & 102 \end{bmatrix} \mathrm{X} \begin{bmatrix} a \\ b_1 \\ b_2 \end{bmatrix}$$

que constituye un producto de dos matrices "conformes"[3] de dimensiones:

$$(10 \times 1) = (10 \times 3) \times (3 \times 1)$$

o bien expresándolo con notación abreviada vectorial,

$$\mathbf{y} = \mathbf{x} \cdot \mathbf{b}\ ,$$

en donde **y**, **x**, y **b** son, respectivamente, cada una de las matrices numéricas y el vector de los parámetros de la primera ecuación matricial anterior.

[3] Es imperioso señalar que en el álgebra matricial el producto de dos matrices exige que ambas (multiplicando y multiplicador) deban ser "conformes" para la multiplicación; así, se tiene que:

$$\mathbf{A\ (\ m, n)\ x\ B(\ p\ ,\ q) = C(\ m, q)}$$

y para que sean conformes para la multiplicación debe cumplirse que: n = p.

h) Calcúlese la matriz inversa $[x^t x]^{-1}$ con los datos del apartado anterior.

SOLUCIÓN.

$$\left[x^t x\right] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 86 & 29 & 61 & 67 & 45 & 9 & 43 & 37 & 17 & 71 \\ 96 & 71 & 120 & 77 & 80 & 89 & 107 & 81 & 65 & 102 \end{bmatrix} \times \begin{bmatrix} 1 & 86 & 96 \\ 1 & 29 & 71 \\ 1 & 61 & 120 \\ 1 & 67 & 77 \\ 1 & 45 & 80 \\ 1 & 9 & 89 \\ 1 & 43 & 107 \\ 1 & 37 & 81 \\ 1 & 17 & 65 \\ 1 & 71 & 102 \end{bmatrix} =$$

$$\text{(aplicando la regla de Binet} - \text{Cauchy)}$$

$$= \begin{bmatrix} 1.1+1.1+...+1.1 & 1.86+1.29+...+1.71 & 1.96+1.71+...+1.102 \\ 86.1+29.1+...+71.1 & 86.86+29.29+...+71.71 & 86.96+29.71+...+71.102 \\ 96.1+71.1+...+102.1 & 96.86+71.29+...+102.71 & 96.96+71.71+...+102.102 \end{bmatrix} =$$

$$= \begin{bmatrix} 10 & 465 & 888 \\ 465 & 27.101 & 43.140 \\ 888 & 43.140 & 81.546 \end{bmatrix}$$

El determinante de esta matriz simétrica toma el valor:

$$\left|x^t x\right| = \begin{vmatrix} 10 & 465 & 888 \\ 465 & 27.101 & 43.140 \\ 888 & 43.140 & 81.546 \end{vmatrix} = 22.099.781.460 + 2 \times 17.813.368.800 -$$

$$-21.370.330.944 - 18.610.596.000 - 17.632.283.850 = 113.308.266 \neq 0$$

, luego la matriz es invertible (puesto que se trata de una matriz regular, no singular). Como $[x^t x]$ es una matriz simétrica, su transpuesta es la misma matriz; la matriz inversa pedida es, siguiendo el método del determinante de la matriz y después de calcular la correspondiente matriz adjunta:

$$[x^t x]^{-1} = \frac{1}{113.308.266}\begin{bmatrix} 348.918.546 & 389.430 & -4.005.588 \\ 389.430 & 26.916 & 18.480 \\ -4.005.588 & 18.480 & 54.785 \end{bmatrix} =$$

$$= \begin{bmatrix} 3'0794 & 0'003437 & -0'03535 \\ 0'003437 & 0'0002375 & -0'0001631 \\ -0'03535 & -0'0001631 & 0'0004835 \end{bmatrix}$$

i) Estímese la ecuación de regresión múltiple lineal del problema que nos ocupa conociendo la matriz inversa $[x^t x]^{-1}$ que se acaba de calcular.

SOLUCIÓN.
La generalización de este tipo de problemas al caso de un número *k* de variables explicativas o independientes (X_1, X_2, …, X_k) debe resolverse empleando notación y cálculo matricial (en este caso, *k* = 2). Existen programas y paquetes informáticos eficientes al respecto.

Consideraremos el caso más general en el que figuran *k* variables explicativas y hagamos figurar, como un término más, la discrepancia d_i en el modelo de regresión lineal múltiple, que puede expresarse matricialmente con la siguiente notación condensada en donde cada una de las letras representa la matriz o vector correspondiente:

$$y = x \cdot b + d$$

Derivando la expresión minimocuadrática respecto a cada uno de los estimadores de regresión, se llegará a un sistema de ecuaciones normales que posee la siguiente expresión matricial abreviada:

$$x^t \cdot x \cdot b = x^t \cdot y$$

El producto de la matriz transpuesta $\mathbf{x^t}$ por la matriz original **x** es siempre una matriz cuadrada, por lo que suponiéndola regular (no singular) y premultiplicando ambos miembros de la última igualdad por la matriz inversa, resultará la expresión matricial que permite estimar los coeficientes de regresión lineal múltiple por el método MC (de los mínimos cuadrados), que es de la forma:

$$\hat{b} = [x^t x]^{-1} x^t y \,,$$

que en nuestro caso toma la expresión:

$$\begin{bmatrix} \hat{a} \\ \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \begin{bmatrix} 3'0794 & 0'003437 & -0'03535 \\ 0'003437 & 0'0002375 & -0'0001631 \\ -0'03535 & -0'0001631 & 0'0004835 \end{bmatrix} x$$

$$x \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 86 & 29 & 61 & 67 & 45 & 9 & 43 & 37 & 17 & 71 \\ 96 & 71 & 120 & 77 & 80 & 89 & 107 & 81 & 65 & 102 \end{bmatrix} x \begin{bmatrix} 27 \\ 7 \\ 24 \\ 20 \\ 8 \\ 5 \\ 17 \\ 9 \\ 3 \\ 35 \end{bmatrix} =$$

$$= \begin{bmatrix} 3'0794 & 0'003437 & -0'03535 \\ 0'003437 & 0'0002375 & -0'0001631 \\ -0'03535 & -0'0001631 & 0'0004835 \end{bmatrix} x \begin{bmatrix} 155 \\ 9.334 \\ 14.907 \end{bmatrix} =$$

$$= \begin{bmatrix} -17'5745 \\ 0'3182 \\ 0'2059 \end{bmatrix} = \begin{bmatrix} \hat{a} \\ \hat{b}_1 \\ \hat{b}_2 \end{bmatrix}$$

Por lo tanto, la ecuación de regresión múltiple lineal pedida es:

$$\mathbf{Y = -17'5745 + 0'3182\ X_1 + 0'2059\ X_2}$$

que, como puede comprobarse, no difiere sensiblemente de la obtenida por el método más tradicional del apartado b) ni tampoco por el más novedoso aquí propuesto del mismo apartado.

j) Estímese la varianza residual de Y sobre X_1 y X_2 del ejercicio anterior tomando como base la expresión matricial:

$$D = \sum d_i^2 = \left[y - x\hat{b} \right]^t \times \left[y - x\hat{b} \right]$$

SOLUCIÓN.

Como resulta que:

$$\left[y - x\hat{b}\right] = \begin{bmatrix} 27 \\ 7 \\ 24 \\ 20 \\ 8 \\ 5 \\ 17 \\ 9 \\ 3 \\ 35 \end{bmatrix} - \left(\begin{bmatrix} 1 & 86 & 96 \\ 1 & 29 & 71 \\ 1 & 61 & 120 \\ 1 & 67 & 77 \\ 1 & 45 & 80 \\ 1 & 9 & 89 \\ 1 & 43 & 107 \\ 1 & 37 & 81 \\ 1 & 17 & 65 \\ 1 & 71 & 102 \end{bmatrix} x \begin{bmatrix} -17'5745 \\ 0'3182 \\ 0'2059 \end{bmatrix} \right) = \begin{bmatrix} 27 - 29'55 \\ 7 - 6'26 \\ 24 - 26'53 \\ 20 - 19'60 \\ 8 - 13'21 \\ 5 - 3'61 \\ 17 - 18'14 \\ 9 - 10'88 \\ 3 - 1'21 \\ 35 - 26'01 \end{bmatrix} = \begin{bmatrix} -2'55 \\ 0'74 \\ -2'53 \\ 0'40 \\ -5'21 \\ 1'39 \\ -1'14 \\ -1'88 \\ 1'79 \\ 8'99 \end{bmatrix},$$

, entonces D tomará el valor:

$$D = \left[-2'55, 0'74, -2'53, 0'40, -5'21, 1'39, -1'14, -1'88, 1'79, 8'89\right] x \begin{bmatrix} -2'55 \\ 0'74 \\ -2'53 \\ 0'40 \\ -5'21 \\ 1'39 \\ -1'14 \\ -1'88 \\ 1'79 \\ 8'99 \end{bmatrix} =$$

$$= (-2'55)^2 + 0'74^2 + ... + 8'99^2 = 131'55 \text{ m}^2$$

Puede comprobarse que el vector $y - x\hat{b}$ toma los valores $d_i = Y_i - T_i$ de la tabla 7 (aunque aquí se han calculado con dos cifras decimales solamente); en el apartado e) obteníamos que $D = \Sigma d_i^2 = 130$ m^2, pero el valor acabado de hallar ahora (131'55 m^2) resulta, sin duda, más exacto.

Con este resultado, la varianza residual es:

$$s^2_{ry,12} = \frac{D}{n} = \frac{131'55}{10} = 13'155 \text{ m}^2,$$

que resulta ligeramente superior a la obtenida en el apartado anterior e), de 13'000 m^2, y el coeficiente de determinación, por tanto, será:

$$R^2_{y,12} = 1 - \frac{s^2_{ry,12}}{s^2_y} = 1 - \frac{13'155}{104'45} = 1 - 0'126 = 0'874 \ ,$$

muy parecido al que se obtuvo en el apartado anterior e), que era de 0'8755. Correlativamente, el coeficiente correspondiente de correlación múltiple lineal sería, en este último caso:

$$R_{y \cdot 12} = \sqrt{0'874} = 0'9349$$

La varianza explicada por la regresión, en definitiva, vendrá dada por la expresión:

$$s^2_{y'} = s^2_y - s^2_{ry \cdot 12} = 104'45 - 13'155 = 91'295 \text{ m}^2$$

con lo que puede comprobarse que:

$$R^2_{y \cdot 12} = \frac{s^2_{y'}}{s^2_y} = \frac{91'295}{104'45} = 0'874 \text{ , c.s.q.d.}$$

4. EJEMPLO 3

Como ilustración de la aplicación de la técnica estadística del análisis de la varianza a este tipo de problemas topográficos, consideremos ahora los datos o coordenadas relativas de la tabla siguiente, obtenidas también de un terreno de gran desnivel en ladera de montaña, en el que al igual que en el caso anterior, para el logro de una mayor facilidad de cálculo, se han ajustado las cifras decimales hasta los enteros. Así:

Tabla 9. Coordenadas de las estacas de la parcela.

Estacas	**Y = Z (m.)**	**X_1 = X (m.)**	**X_2 = Y (m.)**
1	100	100	100
2	106	104	99
3	107	106	110
4	120	111	126
5	110	111	113
6	116	115	103
7	123	120	102
8	133	124	103
9	137	126	98

Al respecto, puede verse la planta curvada del terreno en cuestión, de configuración sensiblemente triangular, con equidistancia vertical de las curvas de nivel de 1’00 m., de la que hemos levantado tres perfiles coincidentes con los lados de dicha figura geométrica plana (verlos en el anexo nº: 5, “Complementos”):

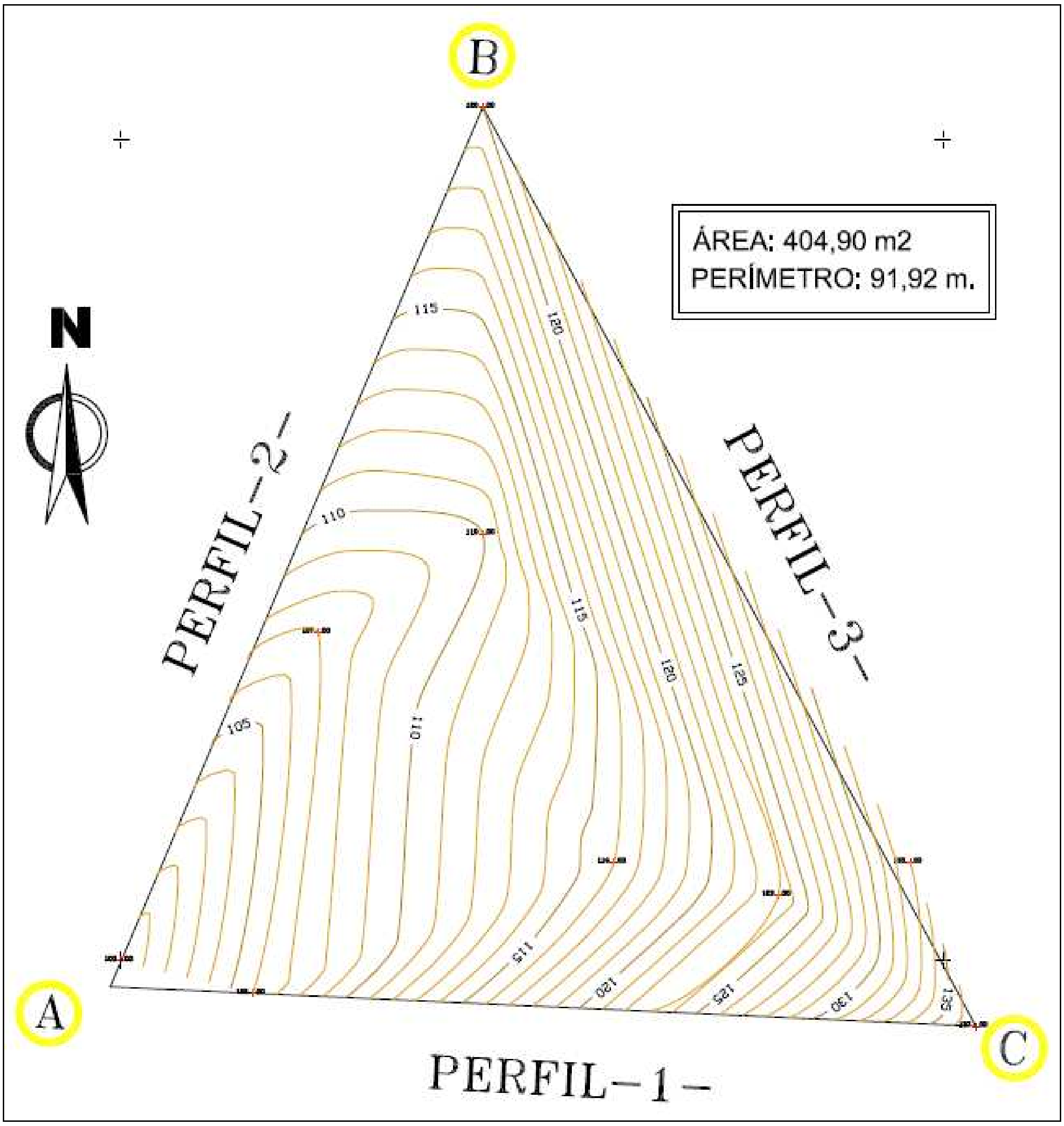

Fig. 4. Planta curvada de la parcela (II).

SOLUCIÓN.

En primer lugar, calculamos los valores para n = 9, que resultan ser los siguientes:

ΣY= 1.052 m.	ΣX_1 = 1.017 m.	ΣX_2 = 954 m.
$\overline{Y}$ = 116’9 m.	$\overline{X}_1$ = 113 m.	$\overline{X}_2$ = 106 m.
ΣY^2 = 124.228 m^2	ΣX_1^2 = 115.571 m^2	ΣX_2^2 = 101.772 m^2
ΣYX_1 = 119.750 m^2	ΣYX_2 = 111.433 m^2	ΣX_1X_2 = 107.690 m^2

Partiendo de aquí construimos las siguientes cantidades en función de las desviaciones con respecto a las medias aritméticas correspondientes, que representaremos con letras minúsculas:

$$\begin{cases} \Sigma y^2 = & 124.228 - \frac{1}{9}(1.052)^2 & = 1.260'89 \text{ m}^2 \\ \Sigma x_1^2 = & 115.571 - \frac{1}{9}(1.017)^2 & = 650 \text{ m}^2 \\ \Sigma x_2^2 = & 101.772 - \frac{1}{9}(954)^2 & = 648 \text{ m}^2 \\ \Sigma yx_1 = & 119.750 - \frac{1}{9}(1.052)(1.017) & = 874 \text{ m}^2 \\ \Sigma yx_2 = & 111.433 - \frac{1}{9}(1.052)(954) & = -79 \text{ m}^2 \\ \Sigma x_1x_2 = & 107.690 - \frac{1}{9}(1.017)(954) & = -112 \text{ m}^2 \end{cases}$$

Tenemos, pues, las siguientes matrices y determinantes:

$$X^tX = \begin{bmatrix} 650 & -112 \\ -112 & 648 \end{bmatrix} \qquad X^tY = \begin{bmatrix} 874 \\ -79 \end{bmatrix}$$

$$\left|X^tX\right| = 408.656 \neq 0, \text{luego la matriz es invertible (regular, no singular), así:}$$

$$(X^tX)^{-1} = \frac{1}{408.656}\begin{bmatrix} 648 & 112 \\ 112 & 650 \end{bmatrix} = \begin{bmatrix} 0'00158568 & 0'00027407 \\ 0'00027407 & 0'00159058 \end{bmatrix}$$

que es la matriz inversa obtenida previo el cálculo de la matriz adjunta transpuesta, esto es, por aplicación del método basado en el cálculo del determinante.

Por consideraciones teóricas acerca de los estimadores mínimo cuadráticos, tenemos que:

$$\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \hat{\beta} = (X^tX)^{-1}X^tY = \begin{bmatrix} 1'36423279 \\ 0'11388140 \end{bmatrix}$$

y del mismo modo:

$$\hat{\beta}_1 = \overline{Y} - \hat{\beta}_2\,\overline{X}_1 - \hat{\beta}_3\,\overline{X}_2 = 116'9 - (1'36423279)(113) - (0'1138814)(106) =$$
$$= -49'3297$$

que da la relación estimada o ecuación del plano de nivelación óptimo ajustando los correspondientes coeficientes hasta las diezmilésimas:

$$\hat{Y} = -49'3297 + 1'3642\ X_1 + 0'1139\ X_2$$

Si ahora, a efectos puramente comparativos, aplicamos el método de cálculo establecido por nosotros a partir de la hoja *Excel* con *Solver*, se tendrá la ecuación que sigue de las cotas definitivas ($Z = T_i$):

$$Y = -49'3333 + 1'3642\ X_1 + 0'1138\ X_2$$

que resulta prácticamente igual que el anterior y que ofrece la siguiente tabla que resulta de mayor precisión:

Tabla 10. Cotas definitivas y correcciones (IV).

Estaca	X_1	X_2	Y_i	T_i	$d_i = Y_i - T_i$
1	100	100	100	98,471	+1,529
2	104	99	106	103,814	+2,186
3	106	110	107	107,795	-0,795
4	111	126	120	116,437	+3,563
5	111	113	110	114,957	-4,957
6	115	103	116	119,276	-3,276
7	120	102	123	125,983	-2,983
8	124	103	133	131,554	+1,446
9	126	98	137	133,713	+3,287
TOTAL	**1.017**	**954**	**1.052**	**1.052,000**	**±0,000**

Para tener una medida objetiva del grado de explanación, en base a lo explicitado en el capítulo 5 anterior, igualaremos en este caso a +20,00 m. la cota relativa media o centro de gravedad de la parcela en estudio, cuyo valor resulta de dividir la suma de las cotas iniciales del terreno natural por el número de vértices (1.052/9 = 116'889 m.), con lo que se tendrá la siguiente tabla:

Vértices	**Cotas relativas iniciales (Y_i)**	**Cotas relativas definitivas (T_i)**	**d_i ($Y_i - T_i$)**	**d_i^2**	**d_i^2/T_i**
1	3,111	1,582	1,529	2,337	1,477
2	9,111	6,925	2,186	4,778	0,690
3	10,111	10,906	-0,795	0,632	0,058
4	23,111	19,548	3,563	12,692	0,649
5	13,111	18,068	-4,957	24,576	1,360
6	19,111	22,387	-3,276	10,731	0,479
7	26,111	29,094	-2,983	8,898	0,306
8	36,111	34,665	1,446	2,092	0,060
9	40,111	36,824	3,287	10,806	0,293
Σ	180,00	180,000	±0,000	77,542	**χ^2=5,373**

Obsérvese que en la tabla anterior hemos definido el $d_i = Y_i - T_i$ como diferencia entre las cotas del terreno natural o iniciales y las definitivas que se deducen de la aplicación de nuestro modelo de explanación. Ello es así con el objetivo de adecuarnos a la terminología utilizada para el cálculo de chi-cuadrado que realizaremos a continuación.

El error estándar o típico de la estima de esta regresión múltiple (triple) vendrá dado por la expresión (véase el capítulo 5) con m = 2 variables explicativas correspondientes a la abscisa y la ordenada de cada punto:

$$S_{xy} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - T_i)^2}{N-m-1}} = \sqrt{\frac{77{,}542}{9-2-1}} = 3'59\ \text{m.}$$

Para N – 1 = 9 – 1 = 8 grados de libertad, se tiene un $\chi^2_{0,5}$ = 1’344, buscando en la tabla de percentiles de la distribución teórica de probabilidad chi-cuadrado que figura en el anexo 3. Al ser: χ^2=5’373 > 1’344 puede considerarse desde luego inaceptable el volumen de explanación a realizar en la parcela que nos ocupa en base al estadígrafo utilizado. Además, en el caso de haber utilizado como centroide o cota relativa media de esta parcela el valor +10,00 m., los valores correspondientes de las cotas T_i que figuran en el denominador hubieran sido menores, con lo que el cociente representado por la χ^2 también hubiera resultado mayor (aproximadamente el doble), o sea, de un valor próximo a 10’7.

Por otra parte, el “grado de explanación” determinado, como ya se ha visto, por el “coeficiente de contingencia” C derivado de la distribución de probabilidad chi-cuadrado (χ^2), vendrá dado por la expresión:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{5'373}{5'373+9}} = 0'61 \cong 61\%,$$

siendo N = 9 el número de estacas o vértices de nivelación considerado.

La suma de cuadrados debida a la influencia lineal de las variables explicativas es:

$$\hat{\beta}^t X^t Y = 1.183'3428$$

lo que ofrece el análisis de la varianza (ANOVA) de la tabla de la página siguiente:

Tabla 11. Análisis de la varianza (I).

Fuente de variación	**Suma de cuadrados**	**Grados de libertad**	**Media**
X_1 y X_2	1.183’34	2	591’67
Residuo	77’55	6	12’93
Total	1.260’89	8	

El valor resultante de F es:

$$F = 591'67/12'93 = 45'76$$

Con (2,6) grados de libertad, se tiene: $F_{0'01} = 10'925$, o bien $F_{0'05} = 5'140$, o bien su valor intermedio: $F_{0'025} = 7'260$, lo que se consigue consultando las tablas correspondientes del anexo 4, de modo que entre estas tres variables o coordenadas del problema existe una asociación altamente significativa.

El análisis puede también realizarse por etapas y analizar la contribución separada de cada variable. En efecto, sea:

$$\begin{cases} b_1 = \text{coeficiente de } X_1 \text{ en la regresión simple de Y respecto a } X_1 \\ b_2 = \text{coeficiente de } X_2 \text{ en la regresión simple de Y respecto a } X_2 \end{cases}$$

Entonces, se cumplirá que:

$$\begin{cases} b_1 = \dfrac{\sum yx_1}{\sum x_1^2} = \dfrac{874}{650} = 1'344615 \\ b_2 = \dfrac{\sum yx_2}{\sum x_2^2} = \dfrac{-79}{648} = -0'121914 \end{cases}$$

La suma explicada de cuadrados debida solamente a la variable X_1 viene dada por:

$$b_1\Sigma yx_1 = (1'344615)\cdot(874) = 1.175'19$$

y, complementariamente, la suma explicada de cuadrados debida solamente a la variable X_2 es:

$$b_2\Sigma yx_2 = (-0'121914)\cdot(-79) = 9'63$$

De estas cantidades así obtenidas podemos conformar las siguientes tablas:

Tabla 12. Análisis de la varianza (II).

Fuente de variación	**Suma de cuadrados**	**Grados de libertad**	**Media**
X_1 Adición de X_2	1.175'19 8'15	1 1	1.175'19 8'15
X_1 y X_2 Residuo	1.183'34 77'55	2 6	 12'93
Total	1.260'89	8	

Tabla 13. Análisis de la varianza (III).

Fuente de variación	**Suma de cuadrados**	**Grados de libertad**	**Media**
X_2 Adición de X_1	9'63 1.173'71	1 1	9'63 1.173'71
X_1 y X_2 Residuo	1.183'34 77'55	2 6	 12'93
Total	1.260'89	8	

La suma total de cuadrados debida a X_1 y a X_2, según la tabla 11, es 1.183'34. En la tabla 12 veamos que la suma de cuadrados de X_1 es 1.175'19 y el efecto *adicional* debido a la inclusión de X_2 lo hallamos por diferencia entre ambos, que es 8'15. El efecto adicional de X_2 se prueba luego mediante la razón F,

$$F = 8'15/12'93 = 0'63$$

con (1,6) grados de libertad, el cual resulta evidentemente no significativo. La significación de sólo X_1, se puede probar calculando la suma de cuadrados de los residuos para X_1, o sea, 1.260'89 – 1.175'19 = 85'70 con 7 grados de libertad, lo que da una media de 12'24. La razón F apropiada es, entonces:

$$F = 1.175'19/12'24 = 96'01$$

con (1,7) grados de libertad, la cual sí que resulta altamente significativa.

Alternativamente, podemos construir la tabla 13, a continuación de la anterior. El efecto directo de la variable X_2 es evidentemente no significativo y, por el contrario, el efecto adicional de la variable X_1 sí es altamente significativo.

El efecto neto (adicional) de las variables del problema planteado X_1 o X_2 podría, complementariamente, haberse probado también utilizando la expresión:

$$t = \frac{\hat{\beta}_i - \beta_i}{\sqrt{\sum_{i=1}^{n} e_i^2 /(n-k)}\sqrt{a_{ii}}}$$

que se distribuye como una *t* de Student (Gosset) con *n-k* grados de libertad, en que a_{ii} es el adecuado elemento diagonal de la matriz inversa $(X^tX)^{-1}$. Para $\hat{\beta}_2$ tenemos, en la hipótesis de cumplirse: $\beta_2 = 0$, que:

$$t = \frac{1'36423279}{\sqrt{12'93} \times \sqrt{0'00158568}} = 9'5275$$

puesto que $\Sigma e^2/(n-k)$ = 12'93, conforme se indica en la tabla 11, y 0'00158568 es el primer elemento y, en su consecuencia, el elemento correspondiente a X_1, en la diagonal principal de la matriz inversa $(X^tX)^{-1}$.

Elevando al cuadrado la expresión anterior, tenemos que:

$$t^2 = 90'77$$

El efecto adicional de X_1 de la tabla 13 viene dado justamente por:

$$F = 1.173'71/12'93 = 90'77$$

y las pruebas son, pues, exactamente equivalentes.

Ahora, calculamos un intervalo de confianza del 95 por ciento para β_2 del siguiente modo, teniendo en cuenta que un intervalo de confianza de un 100(1-e) por ciento para β_2 viene dado por la expresión:

$$\hat{\beta}_i \pm t_{\varepsilon/2}\sqrt{\frac{\sum e_i^2}{n-k}}\sqrt{a_{ii}}$$

esto es:

$$1'3642328 \pm 2'4469\sqrt{12'93} \times \sqrt{0'00158568}$$

es decir:

$$\beta_2 \in (1'0143,\ldots,1'7141)$$

Como una alternativa de lo anteriormente expuesto, veamos que el proceso de cálculo que nos ocupa se puede iniciar partiendo del origen cero y escribiendo la matriz simétrica X^tX como sigue:

$$X^t X = \begin{bmatrix} 9 & 1.017 & 954 \\ 1.017 & 115.571 & 107.690 \\ 954 & 107.690 & 101.772 \end{bmatrix}$$

La desventaja de este nuevo enfoque, por lo menos cuando se trabaja con sencillas calculadoras de bolsillo, es que la primera fila y la primera columna de la matriz anterior contienen, por lo general, elementos mucho más pequeños que el resto de la matriz, por lo que resulta difícil retener suficiente número de cifras significativas en todos los elementos en los cálculos sucesivos a realizar. Por supuesto, el empleo de una hoja de cálculo adecuada por ordenador puede resolver eficazmente el problema planteado.

5. EJEMPLO 4

Sea una parcela o porción de terreno, en la cual se han tomado las coordenadas relativas de los 12 puntos que se expresan en la tabla siguiente y en la que, para mayor simplificación del cálculo, se han obviado las cifras decimales de las coordenadas ajustándolas a las enteras, así:

Tabla 14. Coordenadas de los vértices de la parcela (III).

VÉRTICES	**COORDENADAS RELATIVAS**		
	X = X_1 (m.)	**Y = X_2 (m.)**	**Z = Y (m.)**
1	57	8	64
2	59	10	71
3	49	6	53
4	62	11	67
5	51	8	55
6	50	7	58
7	55	10	77
8	48	9	57
9	52	10	56
10	42	6	51
11	61	12	76
12	57	9	68

A continuación, puede verse el plano altimétrico correspondiente en planta sensiblemente alargada, con sus curvas de nivel dibujadas a una equidistancia vertical de 1’00 m.:

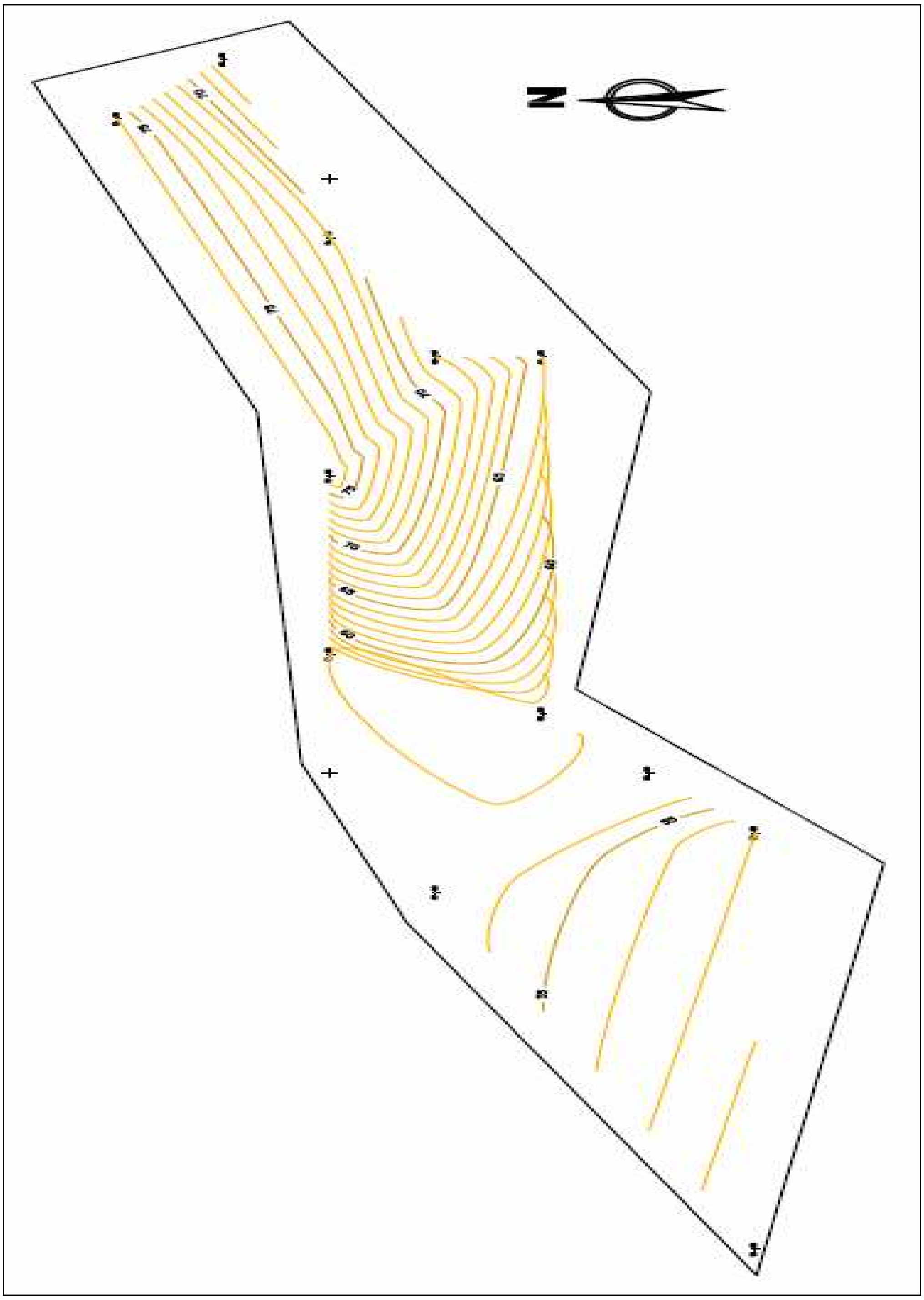

Fig. 5. Planta curvada de la parcela (III).

a) **Estímese la ecuación de regresión lineal múltiple que determina las cotas del plano óptimo definitivo de ajuste por el método tradicional de los mínimos cuadrados.**

SOLUCIÓN.

Puesto que n = 12 vértices o estacas, formaremos la siguiente tabla auxiliar de cálculo:

Tabla 15. Tabla auxiliar de cálculo (IV).

Estaca	Y	X_1	X_2	Y^2	X_1^2	X_2^2	$Y \cdot X_1$	$Y \cdot X_2$	$X_1 \cdot X_2$
1	64	57	8	4.096	3.249	64	3.648	512	456
2	71	59	10	5.041	3.481	100	4.189	710	590
3	53	49	6	2.809	3.401	36	2.597	318	294
4	67	62	11	4.489	3.844	121	4.154	737	682
5	55	51	8	3.025	2.601	64	2.805	440	408
6	58	50	7	3.364	2.500	49	2.900	406	350
7	77	55	10	5.920	3.025	100	4.235	770	550
8	57	48	9	3.249	2.304	81	2.736	513	432
9	56	52	10	3.136	2.704	100	2.912	560	520
10	51	42	6	2.601	1.764	36	2.142	306	252
11	76	61	12	5.776	3.721	144	4.636	912	732
12	68	57	9	4.624	3.249	81	3.876	612	513
TOTAL	**753**	**643**	**106**	**48.139**	**34.843**	**976**	**40.830**	**6.796**	**5.779**

Las ecuaciones normales serán:

$$\begin{cases} \Sigma\, Y = b_{1\cdot 23} \cdot n + b_{12\cdot 3} \cdot \Sigma\, X_1 + b_{13\cdot 2} \cdot \Sigma\, X_2 \\ \Sigma\, Y \cdot X_1 = b_{1\cdot 23} \cdot \Sigma\, X_1 + b_{12\cdot 3} \cdot \Sigma\, X_1^2 + b_{13\cdot 2} \cdot \Sigma\, X_1 \cdot X_2 \\ \Sigma\, Y \cdot X_2 = b_{1\cdot 23} \cdot \Sigma\, X_2 + b_{12\cdot 3} \cdot \Sigma\, X_1 \cdot X_2 + b_{13\cdot 2} \cdot \Sigma\, X_2^2 \end{cases}$$

que, substituyendo los valores obtenidos en la tabla anterior resulta el sistema:

$$\begin{cases} 753 = b_{1\cdot 23} \cdot 12 + b_{12\cdot 3} \cdot 643 + b_{13\cdot 2} \cdot 106 \\ 40.830 = b_{1\cdot 23} \cdot 643 + b_{12\cdot 3} \cdot 34.843 + b_{13\cdot 2} \cdot 5.779 \\ 6.796 = b_{1\cdot 23} \cdot 106 + b_{12\cdot 3} \cdot 5.779 + b_{13\cdot 2} \cdot 976 \end{cases}$$

Resolviendo el anterior sistema de ecuaciones por aplicación de la conocida Regla de Cramer se obtienen los siguientes coeficientes de regresión lineal:

$$b_{1\cdot 23} = 3{,}6512\ ; \quad b_{12\cdot 3} = 0{,}8546\ ; \quad b_{13\cdot 2} = 1{,}5063\ ;$$

con lo que la ecuación de regresión pedida del plano de nivelación adoptará la configuración analítica siguiente:

$$\mathbf{Y = 3{,}6512 + 0{,}8546 \cdot X_1 + 1{,}5063 \cdot X_2}$$

con lo que se da cumplida respuesta a la cuestión planteada.

b) Estímese la misma ecuación de regresión lineal múltiple del plano óptimo de nivelación por aplicación del método presentado en este libro, así como el correspondiente "grado de explanación".

SOLUCIÓN.

Si ahora, a efectos puramente comparativos, aplicamos el método de cálculo establecido por nosotros a partir de la expresada hoja *Excel* con *Solver*, que venimos desarrollando en el presente libro mediante numerosos ejemplos, se tendrá la siguiente ecuación de las cotas definitivas ($Z = T_i$):

$$\mathbf{Y = 3'6540 + 0'8545 \cdot X_1 + 1'5066 \cdot X_2}$$

que resulta más ajustada que la determinación anterior, sin duda con mayor simplicidad de cálculo, y que ofrece la siguiente tabla de discrepancias o correcciones de las cotas taquimétricas inicialmente levantadas:

Tabla 16. Cotas definitivas y correcciones (V).

Estaca	X_1	X_2	Y	T_i	$d_i = Y_i - T_i$
1	57	8	64	64,413	-0,413
2	59	10	71	69,136	+1,864
3	49	6	53	54,564	-1,564
4	62	11	67	73,206	-6,206
5	51	8	55	59,286	-4,286
6	50	7	58	56,925	+1,075
7	55	10	77	65,718	+11,283
8	48	9	57	58,229	-1,229
9	52	10	56	63,154	-7,154
10	42	6	51	48,583	+2,417
11	61	12	76	73,858	+2,142
12	57	9	68	65,920	+2,080
TOTAL	**643**	**106**	**753**	**753,000**	**±0,000**

Para tener una medida objetiva del grado de explanación, en base a lo explicitado en el capítulo 5 anterior, igualaremos a +20,00 m. la cota relativa media o centro de gravedad (centroide) de la parcela en estudio, cuyo valor resulta de dividir la suma de las cotas iniciales del terreno natural por el número de vértices (753/12 = 62'75 m.), con lo que se tendrá la siguiente tabla:

Vértices	Cotas relativas iniciales (Y_i)	Cotas relativas definitivas (T_i)	d_i (Y_i - T_i)	d_i^2	d_i^2/T_i
1	21,25	21,663	-0,413	0,171	0,008
2	28,25	26,386	1,864	3,476	0,132
3	10,25	11,814	-1,564	2,446	0,207
4	24,25	30,456	-6,206	38,509	1,264
5	12,25	16,536	-4,286	18,372	1,111
6	15,25	14,175	1,075	1,155	0,081
7	34,25	22,968	11,283	127,295	5,542
8	14,25	15,479	-1,229	1,511	0,098
9	13,25	20,404	-7,154	51,180	2,508
10	8,25	5,833	2,417	5,844	1,002
11	33,25	31,108	2,142	4,589	0,148
12	25,25	23,170	2,080	4,327	0,187
Σ	240,00	240,000	$\pm$0,000	258,877	**χ^2=12,288**

Obsérvese que en la tabla anterior hemos definido el $d_i = Y_i - T_i$ como diferencia existente entre las cotas del terreno natural o iniciales y las definitivas que se deducen de la aplicación de nuestro modelo de explanación. Ello es así con el objetivo de adecuarnos a la terminología utilizada para el cálculo de chi-cuadrado que realizaremos a continuación.

El error estándar o típico de la estima de esta regresión múltiple (triple) vendrá dado por la expresión (véase el capítulo 5) con m = 2 variables explicativas correspondientes a la abscisa y la ordenada de cada punto:

$$S_{xy} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - T_i)^2}{N - m - 1}} = \sqrt{\frac{258{,}877}{12 - 2 - 1}} = 5'36\,\text{m.}$$

Para N – 1 = 12 – 1 = 11 grados de libertad, se tiene un $\chi^2_{0,5}$ = 2’603, buscando en la tabla de percentiles de la distribución teórica de probabilidad chi-cuadrado que figura en el anexo 3. Al ser: χ^2=12’288 > 2’603 puede considerarse desde luego inaceptable el volumen de explanación a realizar en la parcela que nos ocupa en base al estadígrafo utilizado, al igual que sucedía con la parcela del ejemplo anteriormente desarrollado. De haber empleado una cota media relativa o centroide de +10’00 m., el valor resultante del estadígrafo χ^2 sería aproximadamente el doble, o sea, 24’5.

Por otra parte, el “grado de explanación” determinado, como ya se ha visto, por el “coeficiente de contingencia” C derivado de la distribución de probabilidad chi-cuadrado (χ^2), vendrá dado por la expresión:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{12'288}{12'288 + 12}} = 0'71 \cong 71\%,$$

siendo N = 12 el número de estacas o vértices de nivelación considerado.

Comparando los valores obtenidos de χ^2, S_{xy} y C para esta parcela y la del ejemplo anterior, se observa que en este segundo caso la explanación a efectuar resultará todavía mayor y por tanto muy desfavorable, exigiendo un mayor movimiento de tierras, tanto de desmonte como de terraplén, con el incremento de coste correspondiente.

c) Calcular las desviaciones típicas o *standard* de las tres coordenadas del problema.

SOLUCIÓN.

$$s_1 = \sqrt{\frac{\sum_{i=1}^{12} Y^2}{n} - \left(\frac{\sum_{i=1}^{12} Y}{n}\right)^2} = \sqrt{\frac{48.139}{12} - \left(\frac{753}{12}\right)^2} = 8'6035 \text{ m.} \quad \text{(cota taquimétrica)}$$

$$s_2 = \sqrt{\frac{\sum_{i=1}^{12} X_1^2}{n} - \left(\frac{\sum_{i=1}^{12} X_1}{n}\right)^2} = \sqrt{\frac{34.843}{12} - \left(\frac{643}{12}\right)^2} = 5'6930 \text{ m.} \quad \text{(abscisa)}$$

$$s_3 = \sqrt{\frac{\sum_{i=1}^{12} X_2^2}{n} - \left(\frac{\sum_{i=1}^{12} X_2}{n}\right)^2} = \sqrt{\frac{976}{12} - \left(\frac{106}{12}\right)^2} = 1'8181 \text{ m.} \quad \text{(ordenada)}$$

d) Calcular el coeficiente de correlación múltiple de Y (cota) sobre las restantes coordenadas X_1 (abscisa) y X_2 (ordenada), así como los restantes coeficientes de correlación múltiple entre las otras variables y sus correspondientes coeficientes de determinación o críticos.

SOLUCIÓN.

$$R_{1\cdot 23} = \left[1 - \frac{(64-64'413)^2 + ... + (68-65'920)^2/12}{(64-62'749)^2 + ... + (68-62'749)^2/12}\right]^{1/2} = \left[1 - \frac{21'62}{73'96}\right]^{1/2} = 0'8418$$

Este coeficiente, cuyo valor oscila entre 0 y 1, resulta bastante elevado en nuestro caso e indica la existencia de una correlación grande entre las tres coordenadas del problema planteado, hecho que intuitivamente podríamos haber adelantado. Cuanto más próximo estuviera de 0 la relación lineal sería peor (lo que nos induciría a buscar otros ajustes no lineales como los relacionados en otros apartados de este mismo libro) y cuanto más próximo a 1 sería mejor (si dicho coeficiente alcanza el valor 1 dícese, en tal caso, que la correlación es *perfecta*).

Nótese también que dicho coeficiente de correlación múltiple es mayor que cualquiera de los coeficientes de correlación lineal r_{12} y r_{13}. Esto ocurre siempre y es un hecho que cabía esperar, puesto que se tienen en cuenta variables independientes adicionales adecuadas, llegándose a una relación mejor entre las variables o coordenadas del problema.

Su valor también puede determinarse teniendo en cuenta que:

$$\begin{cases} \text{Variación no explicada} = (64 - 64'413)^2 + ... + (68 - 65'920)^2 = 258'88 \text{ m}^2 \\ \text{Variación total} = 48.139 - 12 \times 62'749^2 = 889'76 \text{ m}^2 \ (258'88 + 630'88) \\ \text{Variación explicada} = 889'76 - 258'88 = 630'88 \text{ m}^2 \end{cases}$$

De aquí se deduce que:

$$R_{1\cdot 23} = \sqrt{630'88/889'76} \equiv 0'8418, \text{ c.s.q.d.}$$

A partir del conocimiento de los coeficientes de correlación lineal entre las variables o coordenadas, tomadas dos a dos, también se puede calcular dicho coeficiente de correlación múltiple aplicando la siguiente fórmula:

$$R_{1\cdot 23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}} =$$

$$= \sqrt{\frac{(0'8196)^2 + (0'7698)^2 - 2(0'8196)(0'7698)(0'7984)}{1 - (0'7984)^2}} = 0'8418$$

cuyo valor, como puede observarse, coincide exactamente con el anteriormente calculado por cualquiera de las formulaciones anteriores.

Su coeficiente de determinación múltiple de Y sobre X_1 y X_2 es:

$$R^2_{1\cdot 23} = 0'8418^2 = 0'7086$$

lo que significa que casi el 71% de la variación total de la cota taquimétrica se explica por medio de la ecuación de regresión hallada.

Del mismo modo, se calcularán los restantes coeficientes de correlación y de determinación múltiple, esto es:

$$R_{2\cdot 13} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1-r_{13}^2}} =$$

$$= \sqrt{\frac{(0'8196)^2 + (0'7984)^2 - 2(0'8196)(0'7698)(0'7984)}{1-(0'7698)^2}} = 0'8606$$

$$R^2_{2\cdot 13} = 0'8606^2 = 0'7406$$

$$R_{3\cdot 12} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1-r_{12}^2}} =$$

$$= \sqrt{\frac{(0'7698)^2 + (0'7984)^2 - 2(0'8196)(0'7698)(0'7984)}{1-(0'8196)^2}} = 0'8234$$

$$R^2_{3\cdot 12} = 0'8234^2 = 0'6780$$

Los resultados anteriores ponen de manifiesto el hecho de que, en general, dichos coeficientes múltiples no tienen por qué ser necesariamente iguales, cuestión ésta que puede demostrarse teóricamente.

e) Calcular los tres coeficientes de correlación parciales lineales.

SOLUCIÓN.

$$r_{12\cdot 3} = \frac{r_{12} - r_{13}\cdot r_{23}}{\left[(1-r_{13}^2)\cdot(1-r_{23}^2)\right]^{1/2}} \qquad (1)$$

Siendo r_{12}, r_{13} y r_{23} los coeficientes de correlación lineal entre las variables. Veamos que se calculan mediante las fórmulas:

$$r_{12} = \frac{n\cdot\sum Y\cdot X_1 - \sum Y\cdot\sum X_1}{\left[(n\cdot\sum Y - (\sum Y)^2)\cdot(n\cdot\sum X_1 - (\sum X_1)^2\right]^{1/2}} \qquad (2)$$

Substituyendo valores tenemos:

$$r_{12} = \frac{12\cdot(40'830) - 753'643}{\left[(12\cdot 48'139 - (753)^2)\cdot(12\cdot 34.843 - (643)^2\right]^{1/2}} = 0'8196$$

Análogamente se hallan r_{13}, poniendo en la fórmula (2) X_2 en lugar de X_1, y nos da un valor de 0'7698 y r_{23} poniendo en la fórmula (2) en lugar de Y, X_2. Y tenemos un valor de r_{23} = 0'7984.

Con estos tres valores entramos en la fórmula (1) y tenemos el coeficiente de correlación parcial entre Y y X_1, considerando X_2 constante y es $r_{12\cdot 3}$ = 0'5334. Del mismo modo calculamos los restantes coeficientes, a saber: $r_{13\cdot 2}$ = 0'3346 y $r_{23\cdot 1}$ = 0'4580.

Disponiendo los coeficientes de correlación ajustados hasta las centésimas en una tabla comparativa veremos lo siguiente:

r_{12} = 0'82	.....	$r_{12\cdot 3}$ = 0'53
r_{13} = 0'77	.....	$r_{13\cdot 2}$ = 0'33
r_{23} = 0'80	.....	$r_{23\cdot 1}$ = 0'46

Observamos que en los tres casos, al obligar a que una variable sea constante la correlación entre las otras dos coordenadas disminuye ostensiblemente.

f) Calcular el error típico de la estima o *standard*.

SOLUCIÓN.

De la tabla anterior se tiene que:

$$s_{1\cdot 23} = \sqrt{\frac{\sum_{i=1}^{12} d_i^2}{n}} = \sqrt{\frac{(-0'413)^2 + 1'864^2 + \ldots + 2'08^2}{12}} = 4'6447 \text{ m.}$$

El error típico de la estima de la población de puntos del terreno que ahora nos ocupa, que posee propiedades análogas a las de la desviación típica o *standard*, viene calculado, como ya hemos visto, por la expresión:

$$\hat{s}_{1\cdot 23} = \sqrt{n/(n-3)}\cdot s_{1\cdot 23} = 5'36 \text{ m.}$$

A similar resultado, sin duda alguna, se habría llegado de haber aplicado, para la resolución del problema planteado, la siguiente fórmula alternativa:

$$s_{1\cdot 23} = s_1 \times \sqrt{\frac{1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}} =$$

$$= 8'6035 \times \sqrt{\frac{1 - 0'8196^2 - 0'7698^2 - 0'7984^2 + 2(0'8196)(0'7698)(0'7984)}{1 - 0'7984^2}} =$$

$$= 4'6447 \text{ m., c.s.q.d.}$$

al que habría que aplicar la corrección correspondiente para obtener el pretendido $\hat{s}_{1\cdot 23}$. En cualquier caso, el valor obtenido con estas formulaciones es el mismo que el deducido en el apartado anterior b), como no podía ser de otra manera.

6. EJEMPLO 5

En la práctica de la explanación de terrenos, con independencia de la búsqueda del plano óptimo de nivelación, suele aparecer con frecuencia la condición de que el plano buscado pase precisamente por un punto previamente determinado a efectos constructivos, urbanísticos o estéticos. Pues bien, para la resolución de este tipo de problemas, o de otros parecidos, puede resultar interesante la consideración del presente ejemplo.

Por el punto de coordenadas (x_0, y_0, z_0) se quiere trazar un plano que forme con los planos coordenados un tetraedro de volumen mínimo. Hallar la ecuación de dicho plano.

SOLUCIÓN.

La ecuación del plano pedido se puede escribir del siguiente modo:

$$\frac{X}{a} + \frac{Y}{b} + \frac{Z}{c} = 1$$

que, como debe ser incidente con (x_0, y_0, z_0), deberá cumplir la condición:

$$\frac{x_0}{a} + \frac{y_0}{b} + \frac{z_0}{c} = 1$$

Considérese que el volumen pedido es:

$$V = \frac{1}{3}\left(\frac{1}{2}ab\right)c = \frac{1}{6}abc$$

donde **a**, **b**, **c** son las variables que se obtienen aplicando el método de los operadores o multiplicadores de Lagrange. Formemos, pues, la función de Lagrange que debemos minimizar con condiciones, a saber:

$$L(a,b,c) = \frac{1}{6}abc + \lambda\left(\frac{x_0}{a} + \frac{y_0}{b} + \frac{z_0}{c} - 1\right)$$

con lo que (condición necesaria o de primer grado):

$$L'_a = \frac{1}{6}bc - \frac{\lambda x_0}{a^2} = 0\,; \quad L'_b = \frac{1}{6}ac - \frac{\lambda y_0}{b^2}\,; \quad L'_c = \frac{1}{6}ab - \frac{\lambda z_0}{c^2}$$

de donde se obtiene:

$$\frac{x_0}{a} = \frac{abc}{6\lambda}\,; \quad \frac{y_0}{b} = \frac{abc}{6\lambda}\,; \quad \frac{z_0}{c} = \frac{abc}{6\lambda}$$

o sea:

$$\frac{x_0}{a} = \frac{y_0}{b} = \frac{z_0}{c}$$

y como

$$\frac{x_0}{a} + \frac{y_0}{b} + \frac{z_0}{c} = 1$$

se tiene

$$\frac{x_0}{a} = \frac{y_0}{b} = \frac{z_0}{c} = \frac{1}{3}$$

de donde se obtiene

$$a = 3x_0, \quad b = 3y_0, \quad c = 3z_0$$

y, por tanto, la ecuación pedida del plano será:

$$\frac{X}{3x_0} + \frac{Y}{3y_0} + \frac{Z}{3z_0} = 1$$

La condición suficiente o de 2º grado implica la formación del determinante funcional hessiano orlado relevante que nos confirma que se trata, efectivamente, de un mínimo. Esta comprobación la proponemos como interesante ejercicio recapitulatorio a nuestros amables lectores.

Del mismo modo, se deduce que el valor del operador de Lagrange es:

$$\lambda = \frac{a^2bc}{6x_0} = \frac{(9x_0^2)\cdot 3y_0\cdot 3z_0}{6x_0} = \frac{81x_0y_0z_0}{6} = \frac{27x_0y_0z_0}{2}$$

Y el volumen pedido será:

$$V = \frac{1}{6}abc = \frac{27x_0y_0z_0}{6} = \frac{9x_0y_0z_0}{2}\ m^3$$

y entonces:

$$\boxed{V = \frac{\lambda}{3}}$$

Si ahora aplicamos los conceptos teóricos anteriormente expresados al ejemplo de la parcela real del capítulo anterior, se desea, v. gr., que el plano de nivelación (o de relleno) en cuestión pase por el punto 1 de coordenadas:

($108 = x_0$, $52 = y_0$, $2'241 = z_0$)

Dicho plano será el siguiente:

$$\frac{X}{\underbrace{3\times 108}_{324}} + \frac{Y}{\underbrace{3\times 52}_{156}} + \frac{Z}{\underbrace{3\times 2'241}_{6'723}} = 1$$

$X + 2'077\cdot Y + 48'193\cdot Z = 324$; o sea:

$$\mathbf{X + 2'077\cdot Y + 48'193\cdot Z - 324 = 0}$$

Los puntos de corte de este plano con los tres ejes coordenados serán:

- <u>Corte con el eje Z</u>:

Cuando $\left.\begin{array}{l} X = 0 \\ Y = 0 \end{array}\right\} \Longrightarrow Z = \frac{324}{48'193} = 6'723\ m.$

Al respecto, puede verse la figura tridimensional de la página siguiente:

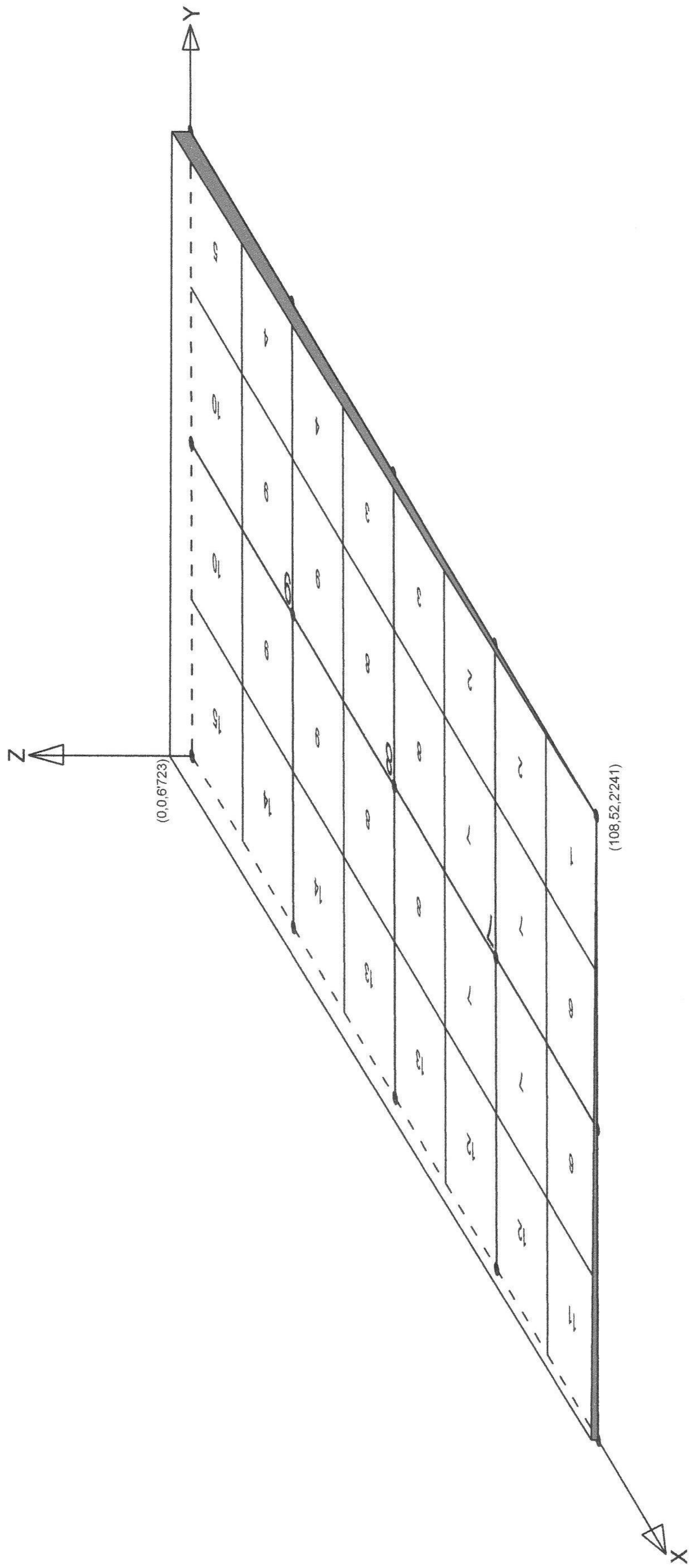

Fig. 6. Tetraedro de volumen mínimo.

- Corte con el eje X:

Cuando $\left.\begin{array}{l} Y = 0 \\ Z = 0 \end{array}\right\} \Longrightarrow X = 324$ m.

- Corte con el eje Y:

Cuando $\left.\begin{array}{l} X = 0 \\ Z = 0 \end{array}\right\} \Longrightarrow Y = \frac{324}{2'077} = 156$ m.

En este caso, el tetraedro completo de relleno tendrá un volumen mínimo de:

$$V = \frac{9 \cdot x_0 \cdot y_0 \cdot z_0}{2} = \frac{9 \times 108 \times 52 \times 2'241}{2} = 56.634'55\ m^3,$$

aunque en el dibujo de la figura anterior se ha considerado únicamente aquella parte del mismo que se proyecta ortogonalmente sobre la parcela o terreno en estudio.

Veamos ahora las coordenadas cartesianas rectangulares de los vértices del cuadrilátero determinado por la malla o red construida sobre la parcela objeto de nuestro estudio. En efecto:

Punto 11

324 m. — 6'723 m.
108 m. — x ; x = 2'241 ; 6'723 – 2'241 = 4'482 m.
$\rightarrow$ **(108,0,4'482)**

Punto 5

156 m. — 6'723 m.
52 m. — x ; x = 2'241 ; 6'723 – 2'241 = 4'482 m.
$\rightarrow$ **(0,52,4'482)**

Punto 1 $\rightarrow$ **(108,52,2'241)**

Punto 15 $\rightarrow$ **(0,0,6'723)**

Cota media de la malla o red: $\frac{4'482 \times 2 + 2'241 + 6'723}{4}$ = **4'482 m.**

que se corresponde exactamente con la cota taquimétrica del centro de gravedad o de masas del terreno en estudio (punto o vértice 8). En efecto, en dicho punto, con X = 54 m. e Y = 26 m., se tiene que:

$$Z = \frac{324 - X - 2'077 \cdot Y}{48'193} = \frac{324 - 54 - 2'077 \cdot 26}{48'193} = \mathbf{4'482\ m.}\text{ , c.s.q.d.}$$

De este modo, la porción del anterior tetraedro comprendida bajo la parcela de terreno que nos ocupa, ocupará un volumen de:

$$V' = 108 \times 52 \times 4'482 = 25.170'91\ m^3 \text{ ,}$$

que supone un:

$$\frac{V'}{V} \times 100 = \frac{25.170'91}{56.634'55} \times 100 = 44'44\%$$

del volumen total del tetraedro mínimo.

Pendientes del bancal:

Las pendientes transversal y longitudinal del plano definitivo de nivelación de la parcela que nos ocupa, en el caso del tetraedro de volumen mínimo, vendrán dadas respectivamente por:

$$\begin{cases} P_t = (2'241/52) \cdot 100 = 4'310\% \\ P_l = (2'241/108) \cdot 100 = 2'075\% \end{cases}$$

Del vértice 1 al 15 existe un desnivel de ΔZ = 4'482 metros por una distancia rectilínea diagonal de:

$$D_{1-15} = \sqrt{108^2 + 52^2} \cong 120\text{ m. ,}$$

con lo que dicha línea tendrá una pendiente de: (4'482/120)·100 = 3'735%.

De este modo, las líneas de máxima pendiente de esta parcela, una vez nivelada, serán:

$$P_m = \sqrt{4'31^2 + 2'075^2} = 4'783\% \text{ (máxima pendiente)}$$

y pueden verse dibujadas en el gráfico siguiente:

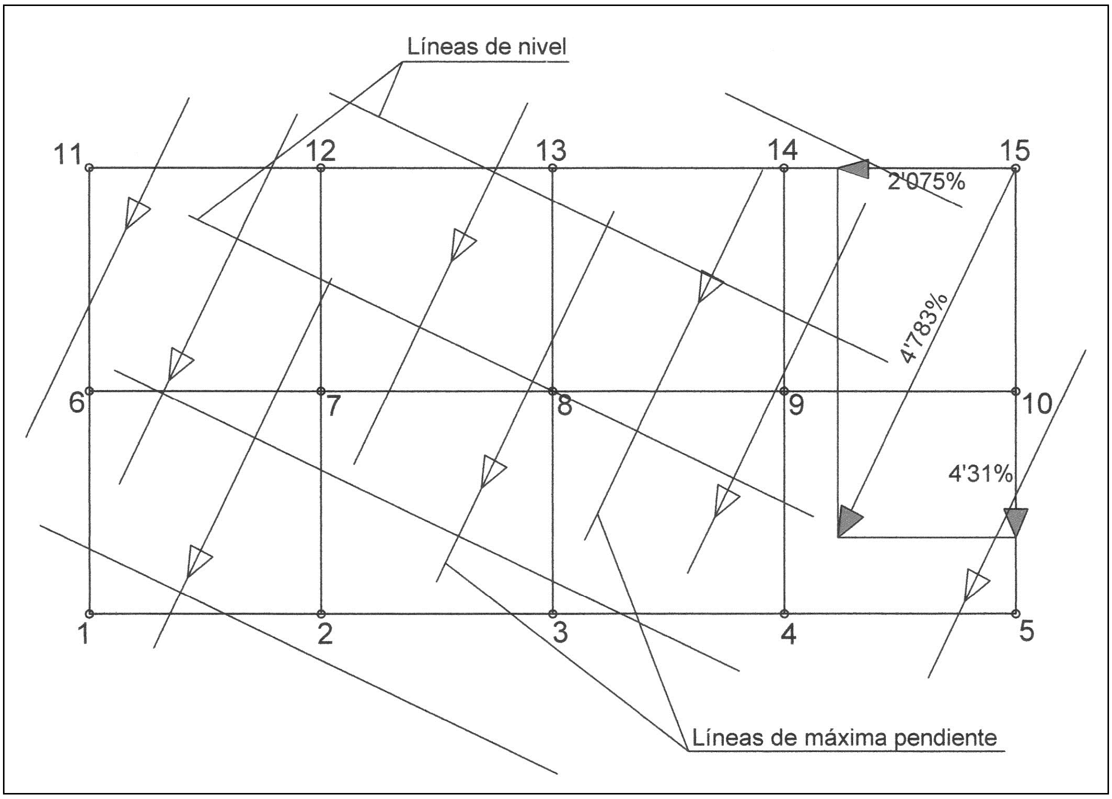

Fig. 7. Líneas de nivel y máxima pendiente de la parcela nivelada (tetraedro de volumen mínimo).

Por cierto que uno de los debates más tradicionales, por lo que se refiere cálculo de la pendiente de un terreno, es si a la hora de medir 1 m. es correcto hacerlo sobre el terreno realmente recorrido, o si se debe tomar 1 m. de avance sobre la horizontal del terreno, o sea, sobre la proyección ortogonal del terreno sobre un mapa.

Tomando el método más purista y exacto, según los topógrafos y geógrafos, se debe tomar el metro recorrido sobre la horizontal, esto es, la base del triángulo que forman la distancia recorrida por el operador (que sería la hipotenusa), la altitud ascendida (que sería el cateto opuesto) y la distancia sobre el mapa (que sería el cateto contiguo). La pendiente es la relación que existe entre el desnivel que debemos superar y la distancia en horizontal que debemos recorrer, lo que equivale a la tangente del ángulo que forma la línea a medir con el eje OX, que sería el plano. La distancia horizontal se mide en el mapa. La pendiente se expresa en tantos por ciento, o bien en grados sexagesimales.

Para calcular una pendiente en tantos por ciento basta con resolver la siguiente regla de tres: “Distancia en horizontal es a 100 como distancia en vertical es a X”, o sea:

Distancia en vertical · 100/Distancia en horizontal = Pendiente (%)

Para calcular la pendiente expresada en grados basta entonces con resolver el triángulo rectángulo con los dos catetos conocidos. Esto es:

Tangente A = Altura/Distancia

Un ángulo de 45º implica una pendiente del 100%, ya que cada 100 metros en horizontal se recorren 100 metros en altura. Cuando medimos una distancia en el mapa lo hacemos sobre una superficie plana. La que medimos en el mapa se llama distancia planimétrica o reducida, que no es otra cosa que la proyección ortogonal en el mapa de la distancia real. La distancia planimétrica coincide con la real sólo si en la realidad hay una llanura, pero si hay una pendiente la diferencia entre la distancia real y la planimétrica puede ser notable.

Para calcular la distancia real debemos hallar el valor de la hipotenusa de un triángulo rectángulo. El valor de un cateto es la distancia en metros entre dos puntos; el valor del otro cateto es el valor en metros de la diferencia en altitud entre los dos puntos del terreno. La distancia real o natural es pues:

$$r^2 = h^2 + a^2 \; ; \; r = \sqrt{h^2 + a^2}$$

Donde:

- **r** = distancia real o natural.
- **h** = distancia horizontal o reducida en la realidad entre los dos puntos.
- **a** = diferencia de altura o desnivel en la realidad entre los dos puntos.

Para medir la distancia existente entre dos puntos del mapa en línea recta basta con usar una regla, un escalímetro o bien nos la proporciona directamente el CAD u otras aplicaciones usuales. Pero en un plano real pocos trazados son rectos. Para medir manualmente trazados sinuosos entre dos puntos se pueden usar dos métodos diferentes: uno rudimentario, que consiste en colocar un hilo sobre el recorrido y luego medir la longitud del hilo; el otro es usando un instrumento creado especialmente para esto, llamado “curvímetro”.

Lo mismo que ocurre con las distancias, una superficie cualquiera del terreno, al venir representada por su proyección ortogonal será en general, salvo en el caso de tratarse de un superficie perfectamente plana y horizontal, de mayor extensión que la expresada en el plano correspondiente. A esta última se la conoce como *superficie agraria* y es la única que hemos de considerar a estos efectos.