Geometría Analítica II

LECTURA 4

Ayudante: Guilmer González 22 de abril, 2014

El día de hoy veremos:

0. Comentarios sobre los trabajos últimos.
1. Un problema de direcciones conjugadas.

1 Direcciones conjugadas

Hemos visto que una forma para encontrar el mínimo de una función

$$f(p) = p^t Ap + 2b^t p + \gamma$$

es calcular direcciones conjugadas

$$d_1,\, d_2,\, d_3$$

de manera que

$$d_1^t Ad_2 = 0, \quad d_2^t Ad_3 = 0, \quad d_3^t Ad_1 = 0$$

La idea en que hemos usado esto es, primero proponer una dirección d_1 cualquiera, luego encontrar d_2 de manera que

$$d_1^t Ad_2 = 0$$

y una vez con ella encontrar d_3 ahora pidiendo

$$d_2^t Ad_3 = 0, \quad d_3^t Ad_1 = 0$$

Una vez que se tengan esos tres vectores o direcciones conjugadas proponemos que el mínimo se logra en

$$p^* = p_0 + \alpha_1^* d_1 + \alpha_2^* d_2 + \alpha_3^* d_3$$

donde usualmente hemos usado $p0 = (0,0,0)^t$ como el punto de partida.

Hemos comprado que el mínimo se logra en el centro de la cónica en $\mathbb{R}^2$ o de la cuádrica si estamos en $\mathbb{R}^3$:

$$Ap^* = -b$$

Lo que hemos estado haciendo se puede interprestar de la siguiente forma:

1) Obtener p_0 de alguna manera (hemos usado el $p_0 = 0$ anteriormente)

2) Proponer d_1 de alguna manera (algunos hemos usado $d_1 = (1,0,0)^t$)

3) Sobre la línea que se forma a partir de p_0 y en la dirección d_1 encontrar el mínimo

$$p_1 = p_0 + \alpha_1^* d_1$$

4) Ahora sobre la línea que se forma entre p_1 y en la dirección d_2 encontrar el mínimo

$$p_2 = p_1 + \alpha_2^* d_2$$

5) Ahora sobre la línea que se forma entre p_2 y en la dirección d_3 encontar el mínimo

$$p_3 = p_2 + \alpha_3^* d_3$$

6) p_3 es el mínimo.

En realidad podría ocurrir que p_2 sea ya el mínimo de f o que lo sea p_1 si tuvieramos tanta suerte al elegir d_1.

Cuando buscamos el mínimo sobre una línea digamos

$$p(\lambda) = q + \lambda d$$

la función sobre esa línea es una cuadrática (para el caso que estamos analizando $A = A^t$ y otras características) para λ

$$
\begin{array}{rcl}
\varphi(\lambda) & = & f(q + \lambda d) \\
 & = & (q + \lambda d)^t A(q + \lambda d) + 2b^t(q + \lambda d) + \gamma \\
 & = & \lambda^2 d^t Ad + 2\lambda(q^t Ad + b^t d) + q^t Aq + 2b^t q + \gamma \\
+ & = & \lambda^2 d^t Ad + 2\lambda(q^t Ad + b^t d) + f(q)
\end{array}
$$

su mínimo se alcanza en el vértice, si la cuadrática se escribe como

$$\varphi(\lambda) = a\lambda^2 + b\lambda + c$$

el vétice de la parábola es $(-b/2a, (b^2 - 4ac)/4a)$, el mínimo se logra en

$$\lambda^* = \frac{-b}{2a}$$

Para el caso que nos compete, haciendo las cuentas necesarias, el mínimo se alcanza en

$$\lambda = \frac{-(q^t Ad + b^t d)}{d^t Ad}$$

Con todo el procedimiento anterior ya podemos calcular el mínimo de f si contamos con la direcciones conjugadas.

El algoritmo que ha visto Pablo en clase es muy particular: propone cómo una forma de calcular directamente esas direcciones. Hagamos un ejemplo que nos permita entender lo que se hace.

Ejemplo : Considere

$$f(x, y) = 6x^2 - 4xy + 9y^2 - 20x - 10y + 1$$

Encontremos el mínimo.

Esta función se puede escribir en forma matricial como

$$f(p) = p^t A p + 2b^t p + \gamma$$

donde

$$A = \begin{pmatrix} 6 & -2 \\ -2 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} -10 \\ -5 \end{pmatrix}, \quad \gamma = 1$$

Usemos el punto $p_0 = (2,3)^t$ como el primer punto y $d_1 = (1,-2)^t$ porque así se nos ha ocurrido.

Calculemos el mínimo sobre esa recta (recuerde que se alcanza sobre el punto medio de la cuerda que cortará a una de las elipses)

$$p_1 = p_0 + \alpha_1 d_1$$

con

$$\alpha_1 = \frac{-(p_0^t A d_1 + b^t d_1)}{d_1^t A d_1}$$

haciendo las cuentas $\alpha_1 = 0.8$ y tenemos que

$$p_1 = p_0 + \alpha_1 d_1 = (2.8, 1.4)^t$$

Debemos calcular d_2 de manera que

$$d_1^A d_2 = 0$$

o bien, mejor escribimos $d_2^t A d_1 = 0$.

Si hacemos la cuentas

$$d_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Ahora calculemos el mínimo sobre

$$p = p_1 + \alpha_2 d_2$$

esto se logra en

$$\alpha_2 = \frac{-(p_1^t A d_2 + b^t d_2)}{d_2^t A d_2}$$

haciendo las cuentas $\alpha_2 = -0.4$ y con esto

$$p_2 = p_1 + \alpha_2 d_2 = (2,1)^t$$

El cual satisface que

$$Ap_2 = -b$$

es el centro de las elipses concéntricas.

En la figura 1 se representa esta idea.

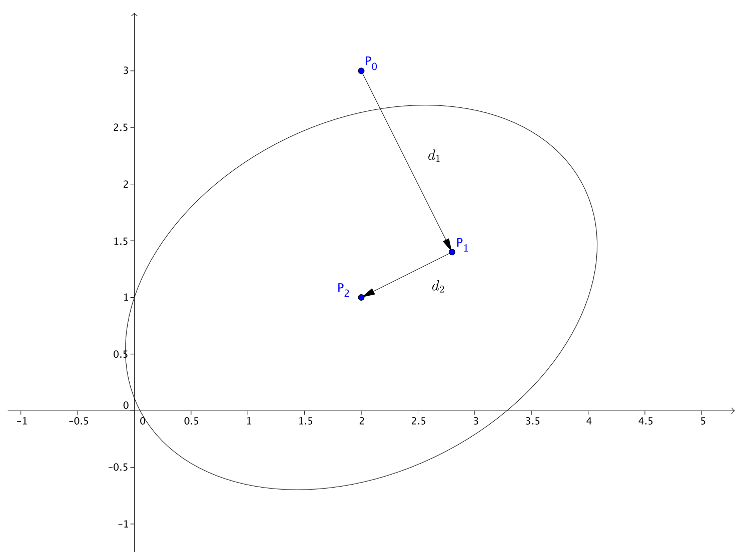

Figura 1: Direcciones conjugadas para calcalcular el mínimo de la función.

Ahora, lo que hemos visto ayer es una forma de calcular las direcciones conjugadas Usemos el mismo ejemplo pero ahora cambiemos a $d_0 = (1, -1)$

Si nosotros hacemos los cálculos obtendemos que el mínimo de f sobre la recta

$$p = p_0 + \alpha d_1$$

se logra en

$$\alpha = \frac{-(p_0^t A d_1 + b^t d_1)}{d_1^t A d_1} = 1.1579$$

y el punto es

$$p_1 = p_0 + \alpha_1 d_1 = [3.1579, 1.8421]$$

pero si evaluamos la función en ese punto observamos que $f(p_1) = -13.4737$, qué pasa si elegimos ahora $d_1 = -g(p_0)$ el vector gradiente de f en p_0? Para la forma en que hemos escrito f, el gradiente se calcula como

$$g(p) = 2Ap + 2b$$

en este caso

$$g(p_0) = 2A * p_0 + 2b = (-8, 36)^t$$

si usamos $d_1 = -g(p_0) = (8, -36)^t$ tenemos que

$$\alpha_1 = -0.0515$$

y que el punto donde se logra el mínimo es

$$p_1 = p_0 + \alpha_1 d_1 = (2.421, 1.1455)^t$$

y el valor de la función en ese punto es

$$f(p_1) = -23.0303$$

La idea es no elegir cualquier vector d_1 en primera instancia, sino elegir $d_1 = -g(p_0)$, como se ve, en algunos casos esa dirección nos podrá ser útil.

Ahora, para el segundo paso, si proponemos d_2 como

$$d_2 = -g_1 + \beta d_1$$

donde

$$g_1 = g(p_1)$$

es decir, una combinación entre d_1 y el gradiente en el punto anterior p_1. Si proponemos esto, el valor de β se puede calcular muy fácilmente (pidiendo que $d_2^t A d_1 = 0$, esto es que sean conjugados):

$$d_2^t A d_1 = 0, \quad \text{ó} \quad (-g_1 + \beta d_1)^t A d_1 = 0$$

y de ahí,

$$\beta = \frac{g_1^t A d_1}{d_1^t A d_1}$$

y listo hemos calculado un vector d_2 conjugado con d_1.

Esto lo que siginifica es que d_2 es una modificación al gradiente g_1 (el gradiente de f en el punto p_1)

Si justificar más, por ahora, y para el caso $\mathbb{R}^3$ vamos a requerir una tercera dirección d_3 conjugada con d_2 y d_1.

Si hacemos

$$d_3 = -g_2 + \beta d_2$$

donde

$$\beta = \frac{g_2^t A d_2}{d_2^t A d_2}$$

Se puede comprobar que si elegimos $d_1 = -g_0$ y las restantes direcciones como se ha dicho, las tres así construidas son conjugadas y esas nos sirven para resolver el problema de encontrar el mínimo de la función f.