UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

Tarea 3

Cálculo III, 525211 (2019)

Catalina Opazo - Patricio Asenjo - Raúl Astete

Problema 1. — Sea (X,dx) un espacio métrico y $A \subset X$ un conjunto compacto.
Demuestre que A es cerrado y acotado.
Solución. Mostraremos en orden que:

1. A es compacto $\Rightarrow$ A es cerrado
2. A es compacto $\Rightarrow$ A es acotado

(1) Probaremos el contrarecíproco, es decir, si A no es cerrado $\Rightarrow$ A no es compacto.

Supongamos que A no es cerrado, es decir $\exists x_0 \in A' : x_0 \notin A$.

$$\begin{aligned} x_0 \in A' \Leftrightarrow &\quad \exists (x^{(n)})_{n\in\mathbb{R}} \in A : \lim_{n\to\infty} x^{(n)} = x_0 \\ \Leftrightarrow &\quad \text{Toda subsucesión } [(x^n)^{(j)}]_{j\in\mathbb{N}} \text{ converge a } x_0 \\ \Leftrightarrow &\quad \exists (x^n)^{(j)} : \lim_{j\to\infty} = x_0 \wedge x_0 \notin A \\ \Rightarrow &\sim (\text{Toda sucesión en A tiene una subsucesión} \\ &\quad\ \text{convergente a un elemento de A}) \\ \Rightarrow &\quad \text{A no es compacto} \end{aligned}$$

(2) Por otro lado veamos que A compacto $\Rightarrow$ A acotado.

Sea A un conjunto compacto, esto es, que para toda familia de conjuntos abiertos que cubre a A, existe una subfamilia **finita** que cubre a A.
Escojamos la familia de bolas centradas en $a \in A$.

$$(B_n)_{n\in\mathbb{N}} = \{B(a,n) : a \in A, \text{ fijo.}\} \text{ tal que } A \subseteq \bigcup_{n\in\mathbb{N}} Bn$$

Como A es compacto, $\exists n \in \mathbb{N} : A \subseteq \bigcup_{i=1}^{n} B_i$. Por lo tanto, si tomamos la bola de centro a y radio n, tenemos que $A \subseteq Bn$, es decir, A esta acotado por esta bola.

Esto es válido, por que la familia $(B_n)_{n\in\mathbb{N}}$ es creciente.

$\therefore$ A compacto $\Rightarrow$ A cerrado y acotado.

Bono: Dé un ejemplo de un conjunto A de elementos de un espacio métrico (X, d_X) que sea cerrado, acotado y no compacto.

Solución. En clases definimos la métrica discreta como

$$\begin{aligned} d_*(\cdot,\cdot) : X &\longrightarrow [0,+\infty[\\ x, y &\longmapsto d_*(x,y) = \begin{cases} 0, \text{si } x = y \\ 1, \text{si } x \neq y \end{cases} \end{aligned}$$

Sea $X = \mathbb{R}$, y sea $A = \{x \in \mathbb{R} : 0 \leq x \leq 1\} \subset X$, tenemos que A es cerrado, pues contiene a todos sus puntos de acumulación, y además, es acotado, pues $\forall x_0 \in A : A \subseteq B(x_0, 1)$. Sin embargo, A no es compacto. En efecto, tenemos que $\forall x_0 \in A : \{x_0\}$ son subconjuntos abiertos de A (por la métrica descrita), y A también posee infinitos elementos. Luego, con el conjunto descrito, no existe recubrimiento finito que cubra a A. Luego, A no es compacto.

Problema 2. — Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ tal que:

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}) + y^2 \sin(\frac{1}{y}) & , \text{si} \quad xy \neq 0 \\ x^2 \sin(\frac{1}{x}) & , \text{si} \quad x \neq 0 \wedge y = 0 \\ y^2 \sin(\frac{1}{y}) & , \text{si} \quad x = 0 \wedge y \neq 0 \\ 0 & , \text{si} \quad x = y = 0 \end{cases} \tag{1}$$

1. Determine si f es continua en R^2.
La función f es continua por ser composicion de funciones continuas $\forall (x,y) \neq (0,0) \wedge (x,y) \neq (x,0) \wedge (x,y) \neq (0,y)$.

Veamos si es continua en (0,0).
Para esto analizaremos $\lim\limits_{(x,y)\to(0,0)} f(x,y) = \lim\limits_{(x,y)\to(0,0)} x^2 \sin\left(\frac{1}{x}\right) + y^2 \sin\left(\frac{1}{y}\right)$

$$\begin{aligned} 0 \leq & \left| x^2 \sin\left(\frac{1}{x}\right) + y^2 \sin\left(\frac{1}{y}\right) \right| \\ \leq & x^2 \left| \sin\left(\frac{1}{x}\right) \right| + y^2 \left| \sin\left(\frac{1}{y}\right) \right| \\ \leq & x^2 + y^2 \xrightarrow{(x,y)\to(0,0)} 0 \end{aligned}$$

Luego, por el Teorema del Sandwich, $\lim\limits_{(x,y)\to(0,0)} x^2 \sin\left(\frac{1}{x}\right) + y^2 \sin\left(\frac{1}{y}\right) = 0 = f(0,0)$. Por lo tanto, f es continua en el $(0,0)$.

Veamos la continuidad de f en puntos de la forma $(x_0, 0)$ y $(0, y_0)$ con $x_0 \neq 0$ y $y_0 \neq 0$.

$$\lim_{(x,y)\to(x_0,0)} x^2 \sin\left(\frac{1}{x}\right) + y^2 \sin\left(\frac{1}{y}\right) = x_0^2 \sin\left(\frac{1}{x_0}\right) = f(x_0, 0)$$

Ya que $\forall \varepsilon > 0 : \exists \delta > 0 : ||(x,y) - (x_0, 0)||_1 < \delta \Rightarrow |f(x,y) - f(x_0, 0)| < \varepsilon$

$$\begin{aligned}
\left|x^2\sin\left(\frac{1}{x}\right)+y^2\sin\left(\frac{1}{y}\right)-x_0^2\sin\left(\frac{1}{x_0}\right)\right| &\leq \left|x^2\sin\left(\frac{1}{x}\right)-x_0^2\sin\left(\frac{1}{x_0}\right)\right|+\left|y^2\sin\left(\frac{1}{y}\right)\right| \\
&\leq \left|x^2\frac{1}{|x|}-x_0^2\frac{1}{|x_0|}\right|+\left|y^2\frac{1}{|y|}\right| \\
&= |x-x_0|+|y| \\
&= ||(x,y)-(x_0,y)||_1 \\
&< \delta
\end{aligned}$$

Escogiendo $\delta=\varepsilon$, mostramos que tal limite existe y que f es continua en los puntos de la forma $(x_0,0)$

Usamos lo mismo para mostrar que

$$\lim_{(x,y)\to(0,y_0)} x^2\sin\left(\frac{1}{x}\right)+y^2\sin\left(\frac{1}{y}\right)=x_0^2\sin\left(\frac{1}{x_0}\right)=f(x_0,0)$$

Y decir que, f es continua en los puntos de la forma $(0,y_0)$.

Finalmente, podemos decir que f es continua en todo $\mathbb{R}^2$.

2. Determine $\partial_1 f$ y $\partial_2 f$. ¿Son estas funciones continuas en $\mathbb{R}^2$?
$\forall(x,y)\neq(0,0)\wedge(x,y)\neq(x_0,0)\wedge(x,y)\neq(0,y_0)$ con $x_0\neq 0, y_0\neq 0$

$$\begin{aligned}
\partial_1 f &= 2x\sin\left(\frac{1}{x}\right)-\cos\left(\frac{1}{x}\right) \\
\partial_2 f &= 2y\sin\left(\frac{1}{y}\right)-\cos\left(\frac{1}{y}\right)
\end{aligned}$$

Para los puntos de la forma $(x,0)$ con $x\neq 0$:

$$\begin{aligned}
\partial_1 f(x,0) &= \lim_{\lambda\to 0}\frac{(x+\lambda,0)-f(x,0)}{\lambda} \\
&= \lim_{\lambda\to 0}\frac{(x+\lambda)^2\sin\left(\frac{1}{x+\lambda}\right)-x^2\sin\left(\frac{1}{x}\right)}{\lambda} \\
&= \lim_{\lambda\to 0}\frac{(x^2+2x\lambda+\lambda^2)\sin\left(\frac{1}{x+\lambda}\right)-x^2\sin\left(\frac{1}{x}\right)}{\lambda} \\
&= \lim_{\lambda\to 0}\frac{x^2\left(\sin\left(\frac{1}{x+\lambda}\right)-\sin\left(\frac{1}{x}\right)\right)}{\lambda}+(2x+\lambda)\sin\left(\frac{1}{x+\lambda}\right) \\
&= \lim_{\lambda\to 0}\frac{x^2\left(\sin\left(\frac{1}{x+\lambda}\right)-\sin\left(\frac{1}{x}\right)\right)}{\lambda}+\lim_{\lambda\to 0}(2x+\lambda)\sin\left(\frac{1}{x+\lambda}\right) \\
&\overset{\text{(L.H)}}{=} \lim_{\lambda\to 0} x^2\ \cos\left(\frac{1}{x+\lambda}\right)\left(\frac{-1}{(x+\lambda)^2}\right)+2x\sin\left(\frac{1}{x}\right) \\
&= -\cos\left(\frac{1}{x}\right)+2x\sin\left(\frac{1}{x}\right)
\end{aligned}$$

$$
\begin{aligned}
\partial_1 f(0,y) &= \lim_{\lambda\to 0} \frac{f(\lambda,y)-f(0,y)}{\lambda} \\
&= \lim_{\lambda\to 0} \frac{\lambda^2 \sin\left(\dfrac{1}{\lambda}\right) + y^2 \sin\left(\dfrac{1}{y}\right) - y^2 \sin\left(\dfrac{1}{y}\right)}{\lambda} \\
&= \lim_{\lambda\to 0} \lambda \sin\left(\frac{1}{\lambda}\right) && \textit{(**)} \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\partial_1 f(0,0) &= \lim_{\lambda\to 0} \frac{f(\lambda,0)-f(0,0)}{\lambda} \\
&= \lim_{\lambda\to 0} \frac{\lambda^2 \sin\left(\dfrac{1}{\lambda}\right) - 0}{\lambda} \\
&= \lim_{\lambda\to 0} \lambda \sin\left(\frac{1}{\lambda}\right) && \textit{(**)} \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\partial_2 f(0,0) &= \lim_{\lambda\to 0} \frac{f(0,\lambda)-f(0,0)}{\lambda} \\
&= \lim_{\lambda\to 0} \frac{\lambda^2 \sin\left(\dfrac{1}{\lambda}\right) - 0}{\lambda} \\
&= \lim_{\lambda\to 0} \lambda \sin\left(\frac{1}{\lambda}\right) && \textit{(**)} \\
&= 0
\end{aligned}
$$

(**) Notar que $0 \leq \left|\lambda \sin\left(\dfrac{1}{\lambda}\right)\right| \leq |\lambda| \xrightarrow{\lambda\to 0} 0$ Asi, por Teorema del Sandwich tal límite es 0.

Calcular $\partial_2 f(0,y)$ es análogo a calcular $\partial_1 f(x,0)$ y calcular $\partial_2 f(x,0)$ es análogo a calcular $\partial_1 f(0,y)$ llegando al resultado

$$
\begin{aligned}
&\partial_2 f(0,y) = -\cos\left(\frac{1}{y}\right) + 2y \sin\left(\frac{1}{y}\right) \\
&\partial_2 f(x,0) = 0
\end{aligned}
$$

Finalmente las derivadas parciales son:

$$
\partial_1 f = \begin{cases} 2x \sin\left(\dfrac{1}{x}\right) - \cos\left(\dfrac{1}{x}\right) & ,\text{si } (x,y) \neq (0,0) \vee (x \neq 0 \wedge y = 0) \\ 0 & ,\text{si } (x = 0 \wedge y = 0) \vee (x = 0 \wedge y \neq 0) \end{cases}
$$

$$\partial_2 f = \begin{cases} 2y\sin\left(\dfrac{1}{y}\right) - \cos\left(\dfrac{1}{y}\right) & ,\text{si } (x,y)\neq(0,0)\vee(x=0\wedge y\neq 0) \\ 0 & ,\text{si } (x=0\wedge y=0)\vee(x\neq 0\wedge x\neq 0) \end{cases}$$

Luego, vemos que ninguna es continua cuando (x,y) tiende a $(0,0)$, $(0,y)$ o $(x,0)$, pues todas son equivalentes a $\lim\limits_{x\to 0}\ 2x\ \sin\left(\dfrac{1}{x}\right) - \cos\left(\dfrac{1}{x}\right)$ y este no existe, pues la función $\cos\left(\dfrac{1}{x}\right)$ diverge cuando x tiende a 0.

3. Determine si f es diferenciable en $(0,0)$
Para aquello, veremos el si el siguiente límite resulta ser 0. Considerando $\nabla f(0,0)=(0,0)$, calculado anteriormente.

$$\begin{aligned} &\lim_{(x,y)\to(0,0)} \frac{|f(x,y)-f(0,0)-\langle\nabla f;(x,y)\rangle|}{||(x,y)||} \qquad (1) \\ \Leftrightarrow &\lim_{(x,y)\to(0,0)} \frac{|x^2\sin(\frac{1}{x})+y^2\sin(\frac{1}{y})|}{\sqrt{x^2+y^2}} \end{aligned}$$

Luego, notemos que

$$\begin{aligned} 0 &\leq \frac{|x^2\sin(\frac{1}{x})+y^2\sin(\frac{1}{y})|}{\sqrt{x^2+y^2}} \\ &\leq \frac{|x^2\sin(\frac{1}{x})|}{\sqrt{x^2+y^2}} + \frac{|y^2\sin(\frac{1}{y})|}{\sqrt{x^2+y^2}} \\ &\leq \frac{x^2}{\sqrt{x^2+y^2}} + \frac{y^2}{\sqrt{x^2+y^2}} \\ &\leq \frac{x^2+y^2}{\sqrt{x^2+y^2}} + \frac{y^2+x^2}{\sqrt{x^2+y^2}} \\ &= 2\sqrt{x^2+y^2} \\ &\xrightarrow{(x,y)\to(0,0)} 0 \end{aligned}$$

Así por el Teorema del Sandwich el límite (1) es 0 y la función es diferenciable en $(x,y)=(0,0)$.

Problema 3. — Sea $p:\mathbb{R}^n\longrightarrow\mathbb{R}$ tal que $\forall x\in\mathbb{R}^n : p(x)=\langle x;x\rangle$.

(3.1) Sea $x_0\in\mathbb{R}^n$. Demuestre, utilizando la definición de diferenciabilidad, que p es diferenciable en x_0 y $\mathrm{d}p(x_0;h)=2\langle x_0;h\rangle$

Solución. Antes de proceder, recordemos algunas propiedades útiles del producto interior y normas en $\mathbb{R}^n$:

- Todas las normas son equivalentes.

- $\|x\| = 0 \Leftrightarrow x = \theta$
- El producto interior es conmutativo.
- El producto interior distribuye en la adición por la izquierda y por la derecha.
- Todo producto interior induce una norma ($\forall x \in \mathbb{R}^n : \sqrt{\langle x; x\rangle} = \|x\|$)

Con esto, procederemos a demostrar que la función es diferenciable en $x_0 \in \mathbb{R}^n$. Por definición:

$$\begin{aligned}
&\lim_{h\to\theta} \frac{|p(x_0+h) - p(x_0) - dp(x_0;h)|}{\|h\|} \\
&= \lim_{h\to\theta} \frac{|\langle x_0+h; x_0+h\rangle - \langle x_0; x_0\rangle - 2\langle x_0; h\rangle|}{\|h\|} \\
&\overset{(\text{Usando propiedades de } \langle\cdot;\cdot\rangle)}{=} \lim_{h\to\theta} \frac{|\langle x_0; x_0\rangle + \langle x_0; h\rangle + \langle h; x_0\rangle + \langle h; h\rangle - \langle x_0; x_0\rangle - 2\langle x_0; h|\rangle}{\|h\|} \\
&\overset{(\text{Como el p.i. es conmutativo en } \mathbb{R}^n)}{=} \lim_{h\to\theta} \frac{|\langle h; h\rangle|}{\|h\|} \\
&= \lim_{h\to\theta} \frac{\|h\|^2}{\|h\|} \\
&= \lim_{h\to\theta} \|h\| = 0.
\end{aligned}$$

Así, queda demostrado lo enunciado.

(3.2) Sean $A \in \mathcal{M}_n(\mathbb{R})$, invertible, $b \in \mathbb{R}^n$ y $r : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ la función que a cada vector de $\mathbb{R}^n$ hace corresponder $r(x) = Ax - b$.

Considere $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que $f(x) = (p \circ r)(x)$. Demuestre, utilizando la regla de la cadena, que

$$df(x_0; h) = 2\langle Ax_0 - b; Ah\rangle.$$

Solución. Sabemos que para cualquier vector de $\mathbb{R}^n$, la función p es diferenciable. Ahora, veamos la diferenciabilidad de r. Sea $y_0 \in \mathbb{R}^n$. Supongamos que es diferenciable en y_0, y que $\mathrm{d}r(y_0; h) = Ah$. Veamos; por definición de diferencial:

$$\begin{aligned}
&\lim_{h\to\theta} \frac{\|r(y_0+h) - r_(y_0) - dr(y_0;h)\|}{\|h\|} \\
&= \lim_{h\to\theta} \frac{\|A(y_0+h) - b - Ay_0 + b - Ah\|}{\|h\|} \\
&= \lim_{h\to\theta} \frac{\|Ay_0 + Ah - Ay_0 - Ah\|}{\|h\|} = 0
\end{aligned}$$

Así, mostramos que r es diferenciable en $y_0 \in \mathbb{R}^n$, con $dr(y_0; h) = Ah$.
Como p y r son funciones diferenciables en todo $\mathbb{R}^n$, entonces es posible aplicar la regla de la cadena para calcular su diferencial. De esta forma

$$df(x_0; h) = d(p \circ r)(x_0; h) = dp(r(x_0); dr(x_0; h)) = 2\langle Ax_0 - b; Ah\rangle.$$

(3.3) El **método del máximo descenso** es un método iterativo para encontrar una aproximación a mínimos de funciones.

En esta tarea derivaremos una versión de este método para la solución de sistemas de ecuaciones lineales que se basa en la propiedad de que si A es una matriz invertible, el vector $x \in \mathbb{R}^n$ que satisface $Ax = b$ también es el que minimiza la función $f(x)$ definida antes.

Sea x_0 una aproximación a la solución exacta de $Ax = b$. Considere la función

$$g(\lambda) = f(x_0 - \lambda d_0)$$

(3.4) Demuestre, con ayuda de la regla de la cadena, que

$$dg(\lambda; s) = 2s\langle Ax_0 - b - \lambda Ad_0; -Ad_0\rangle = 2s(\lambda\|Ad_0\|^2 - \langle Ax_0 - b; Ad_0\rangle)$$

Solución. Definimos la función

$$\begin{aligned} m : \mathbb{R} &\longrightarrow \mathbb{R}^n \\ \lambda &\longmapsto m(\lambda) = x_0 - \lambda d_0 \quad x_0, d_0 \in \mathbb{R}^n \end{aligned}$$

De esta forma, la función g se define como la compuesta de las funciones m y f $(g = f \circ m)$. Antes de seguir, definiremos la diferencial de la función m, la cual está dada por la expresión $\mathrm{d}m(\lambda; s) = -sd_0$, con d_0 un vector de $\mathbb{R}^n$. En efecto:

$$\lim_{s\to 0} \frac{\|m(\lambda + s) - m(\lambda) - dm(\lambda; s)\|}{|s|} = \lim_{s\to 0} \frac{\|x_0 - (\lambda + s)d_0 - x_0 + \lambda d_0 + sd_0\|}{|s|} = 0$$

Con este resultado en cuenta, procederemos a calcular la diferencial de g (Notar que cumple con la hipótesis de la regla de la cadena, pues, al tomar vectores y valores arbitrarios, no se impusieron restricciones sobre la elección de éstos, de lo que se concluye que son diferenciables en todo su dominio):

$$\begin{aligned} dg(\lambda; s) = d(f \circ m)(\lambda; s) = df(m(\lambda); dm(\lambda; s)) &= 2 \cdot \langle A \cdot m(\lambda) - b; A \cdot dm(\lambda; s)\rangle \\ &= 2 \cdot \langle A(x_0 - \lambda d_0) - b; -sAd_0\rangle \\ &= 2s \cdot \langle Ax_0 - b - \lambda Ad_0; -Ad_0\rangle \\ &= 2s \cdot (\langle Ax_0 - b; -Ad_0\rangle + \langle -\lambda Ad_0; -Ad_0\rangle) \\ &= 2s \cdot (-\langle Ax_0 - b; Ad_0\rangle + \lambda\|Ad_0\|^2) \end{aligned}$$

(3.5) Determine el valor de $\lambda_* \in \mathbb{R}$ que minimiza la función g.

Solución. Debemos encontrar el valor de λ_* tal que

$$f(x_0 - \lambda d_0) \leq f(x_0) \qquad \text{(Por observación hecha en el enunciado original)}$$

para así, minimizar la función g. Procediendo a trabajar la desigualdad, se tiene que:

$$\begin{aligned} g(\lambda) &= f(x_0 - \lambda_* d_0) \\ &= p(A(x_0 - \lambda d_0) - b) \\ &= \big\langle Ax_0 - \lambda_* Ad_0 - b; Ax_0 - \lambda_* Ad_0 - b\big\rangle \\ &= \big\langle Ax_0 - \lambda_* Ad_0 - b; -\lambda_* Ad_0\big\rangle + \underbrace{\big\langle Ax_0 - b; Ax_0 - b\big\rangle}_{f(x_0)} + \big\langle -\lambda_* Ad_0; Ax_0 - b\big\rangle \leq f(x_0) \end{aligned}$$

De esta forma, debemos minimizar la expresión

$$\Big\langle Ax_0 - \lambda_* Ad_0 - b; -\lambda_* Ad_0\Big\rangle + \Big\langle -\lambda_* Ad_0; Ax_0 - b\Big\rangle =: k(\lambda)$$

Trabajando esta última función, y buscando su mínimo, tenemos lo siguiente:

$$\begin{aligned} k(\lambda) &= \Big\langle Ax_0 - \lambda_* Ad_0 - b; -\lambda_* Ad_0\Big\rangle + \Big\langle Ax_0 - b; -\lambda_* Ad_0\Big\rangle \\ &= \Big\langle 2Ax_0 - 2b - \lambda_* Ad_0; -\lambda_* Ad_0\Big\rangle \\ &= \langle 2Ax_0 - 2b; -\lambda_* Ad_0\rangle + \langle -\lambda_* Ad_0; -\lambda_* Ad_0\rangle \\ &= -2\lambda\langle Ax_0 - b; Ad_0\rangle + \lambda_*^2\langle Ad_0; Ad_0\rangle \end{aligned}$$

Luego, podemos derivar con respecto a λ_* e igualamos a cero, lo que nos otorga el mínimo de la función, el cual nos da:

$$\lambda_* = \frac{1}{\|Ad_0\|}\Bigg\langle Ax_0 - b; \frac{1}{\|Ad_0\|}Ad_0\Bigg\rangle$$

(3.6) Del proyecto `calculo3_S1_2019` en github baje los siguientes archivos matlab:

- `tarea3_funciongradiente.m`: Archivo matlab con el que se puede graficar una función de $\mathbb{R}^2$ en $\mathbb{R}$, sus curvas de nivel t los vectores gradientes de distintos puntos del dominio.
- `tarea3_maximodescenso.m`: Archivo matlab para utilizar el metodo de máximo descenso para resolver sistemas de ecuaciones lineales.
- Corra el primero de los archivos anteriores con distintas funciones, Observe cómo en cada caso los vectores gradientes sol otrogonales a las curvas de nivel. Explique por qué ocurre esto.

Solucion: Primero veremos por qué el gradiente es ortogonal a las curvas de nivel de una función.
Sea $x_0 \in \mathbb{R}^n$ y $f : \mathbb{R}^n \to \mathbb{R}$. Las derivadas direccionales de f con respecto a un vector $v \in \mathbb{R}^n$ están dadas por

$$\frac{\partial f}{\partial v}(x_0) = \langle \nabla f(x_0)\ ;\ v\rangle$$

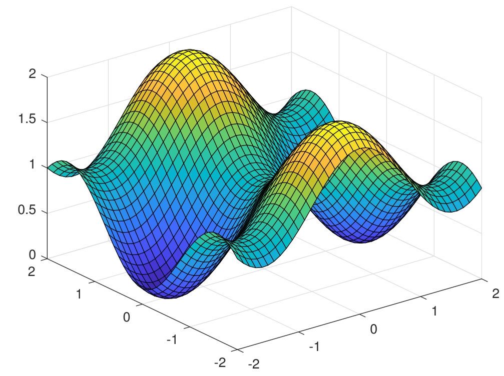
Consideremos ahora v en la dirección de una cuva de nivel $L_c(f)$, como a lo largo de cualquier curva de nivel, el valor de f es constante, la derivada direccional

$$\frac{\partial f}{\partial v}(x_0) = 0$$

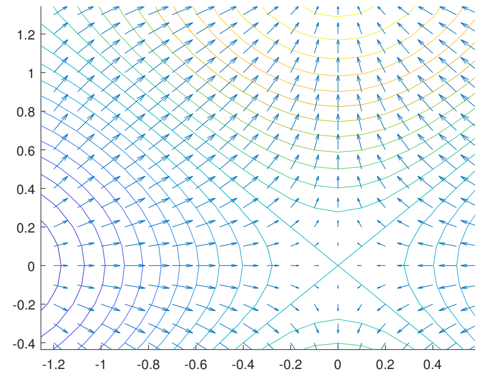
Equivalentemente,

$$\langle \nabla f(x_0)\ ;\ v\rangle = 0$$

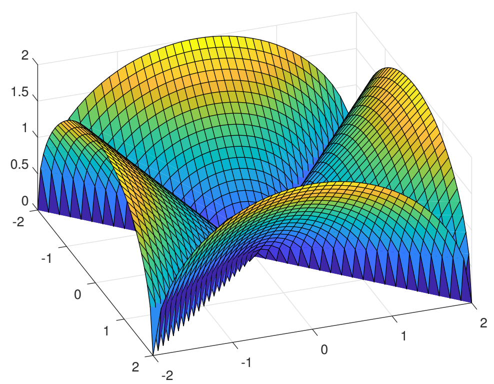
Es decir, el gradiente de f es ortogonal a cualquier vector de la curva de nivel.
Esto lo podemos evidenciar tras graficar varias funciones, pues el resultado que entregan es el siguiente:



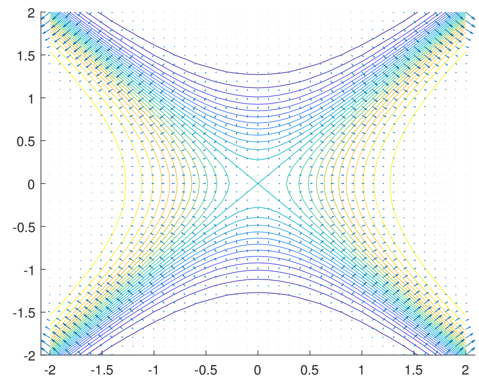
(a) Superficie función 1



(c) Curvas de nivel función 1



(b) Superficie función 2



(d) Curvas de nivel función 2

En cada curva de nivel que logramos graficar, es posible apreciar que las flechas que indican al gradiente son ortogonales a cada curva de nivel.

(3.7) Siga las instrucciones del segundo archivo y encuentre una aproximación a la solución exacta al sistema de ecuaciónes lineales en él con el método del máximo descenso. ¿Cuál es la distancia entre la solución exacta y la aproximación obtenida después de 6 iteraciones?
¿Cuál es la norma de $\|Ax^{(6)} - b\|$ si $x^{(6)}$ representa la aproximación a la solución exacta del sistema de ecuaciones después de 6 iteraciones del método de máximo descenso?

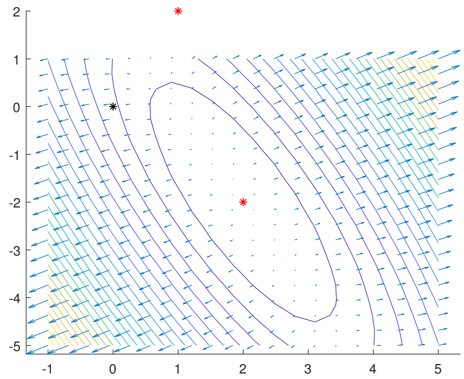
Solución: Se nos presenta el sistema

$$\underbrace{\begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 2 \\ -8 \end{pmatrix}}_{b} \tag{2}$$

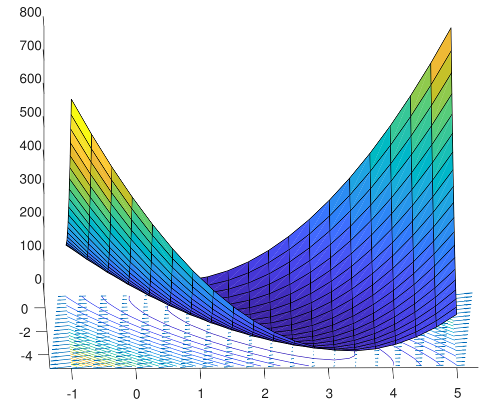
Luego, trabajamos con la función

$$f(z) = \langle Az - b\,;\, Az - b \rangle = \|Az - b\|^2$$

Al graficar, se obtiene



(a) C.de Nivel de f



(b) Superficie y curvas de nivel de $f(z)$

Ahora, nos interesa minimizar la función f, para así encontrar una aproximación de la solución al sistema definido en 2. Debemos recordar que $\|Az - b\|$ define una forma cuadrática.
El método se basa en el siguiente teorema

> **Teorema.** Sea $A \in \mathcal{M}_n(\mathbb{R})$ simétrica y definida positiva, $b \in \mathbb{R}^n$, y J la forma cuadrática definida como
> $$J(x) = \frac{1}{2}x^T Ax - x^T b$$
> Entonces,
> $$\forall y \neq x, \qquad Ax = b \iff J(x) < J(y)$$

Esto nos dice que podemos encontrar la solución al sistema $Ax = b$ minimizando la función $J(x)$.
Lo primero que corresponde analizar es el gradiente de J, si definimos $A = (a_{ij})$ se tiene que

$$J(x) = \frac{1}{2}x^T Ax - x^T b = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}x_i x_j - \sum_{i=1}^{n} b_i x_i$$

Luego,

$$\begin{aligned} \frac{\partial J}{\partial x_k} = \sum_{i=1}^{n} a_{ki}x_i - b_k x_k \quad \implies \quad \nabla J(x) &= \left(\frac{\partial J(x)}{\partial x_1}, \ldots, \frac{\partial J(x)}{\partial x_n}\right) \\ &= \left(\sum_{i=1}^{n} a_{1i}x_i - b_1 x_1, \ldots, \sum_{i=1}^{n} a_{ni}x_i - b_n x_1\right) \\ &= Ax - b \end{aligned}$$

De aquí, obtenemos que la dirección de máximo descenso es $-\nabla J(x) = b - Ax$, a este vector lo llamaremos *residuo* del sistema $Ax = b$, y lo denotaremos con la letra $\mathbf{r}$.

En el problema, la matriz A y el vector b son conocidos. Además, se nos entrega el vector $x^{(0)}$. Supongamos que nuestro algoritmo parte en $x^{(0)}$, el primer paso en la dirección del máximo descenso será el punto

$$x^{(1)} = x^{(0)} + \alpha r^{(0)}$$

Pero el problema es que no conocemos el valor de α, es decir, no sabemos que tan grande debe ser el salto.
La ecuación $x = x^{(0)} + \alpha r^{(0)}$ representa una recta que pasa por el punto $x^{(0)}$ en la dirección de $r^{(0)}$. El problema que tenemos es el de encontrar el mínimo de $J(x)$ en esa recta, podríamos aplicar multiplicadores de Lagrange, pero también podemos encontrar ese mínimo mediante lo que se conoce como una *Line search*, que consiste encontrar el mínimo de la función $f(\alpha) := J(x_0 + \alpha r^{(0)})$, que es una función de una variable real α. Desarrollamos la expresión

$$\begin{aligned}
f(\alpha) &= J(x^{(0)} + \alpha r^{(0)}) \\
&= \frac{1}{2}(x^{(0)} + \alpha r^{(0)})^t A(x^{(0)} + \alpha r^{(0)}) - (x^{(0)} + \alpha r^{(0)})^t b \\
&= \frac{1}{2}({x^{(0)}}^t A x^{(0)} + \alpha {x^{(0)}}^t A r^{(0)} + \alpha {r^{(0)}}^t A x^{(0)} + \alpha^2 {r^{(0)}}^t A r^{(0)}) - {x^{(0)}}^t b - \alpha {r^{(0)}}^t b \\
&= \frac{1}{2}({x^{(0)}}^t A x^{(0)} + 2\alpha {r^{(0)}}^t A x^{(0)} + \alpha^2 {r^{(0)}}^t A r^{(0)}) - {x^{(0)}}^t b - \alpha {r^{(0)}}^t b
\end{aligned}$$

Así,

$$\frac{\partial f(\alpha)}{\partial \alpha} = {r^{(0)}}^t A x^{(0)} + \alpha {r^{(0)}}^t A r^{(0)} - {r^{(0)}}^t b \implies \alpha = \frac{{r^{(0)}}^t (b - A x^{(0)})}{{r^{(0)}}^t A r^{(0)}} = \frac{{r^{(0)}}^t r^{(0)}}{{r^{(0)}}^t A r^{(0)}}$$

Con esto podemos concluir que el método se describe con las ecuaciones

$$\boxed{\begin{aligned}
r^{(k)} &= b - A x^{(k)} \\
\alpha_k &= \frac{{r^{(k)}}^t r^{(k)}}{{r^{(k)}}^t A r^{(k)}} \\
x^{(k+1)} &= x^{(k)} + \alpha_k r^{(k)}
\end{aligned}}$$

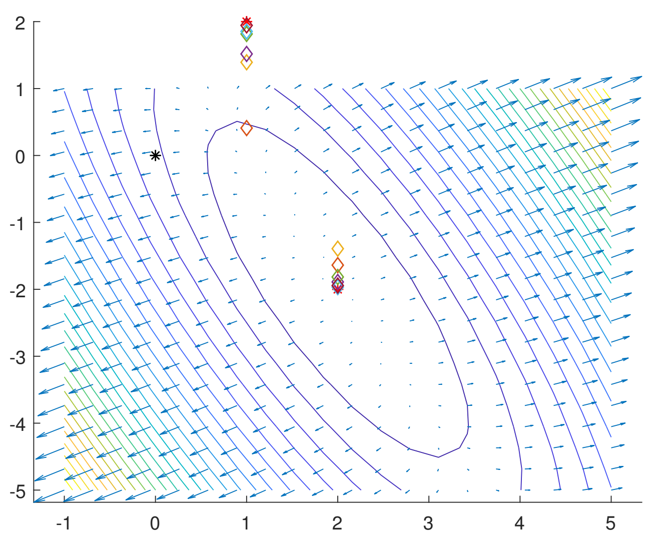
Expresando la fórmula en Matlab, desarrollamos el siguiente código

Listing 1: Programa 1

```
clc;
A = [3 2;2 6];
b = [2;-8];

N = 20;
x = linspace(-1,5,N);
y = linspace(-5,1,N);
[X,Y] = meshgrid(x,y);

Primera Parte: Grafico de f(z)
f = zeros(1,N);

for j=1:N
    for i=1:N
        f(i,j) = dot(A*[x(1,i);y(1,j)]-b,A*[x(1,i);y(1,j)]-b);
    end
end

```

```
surf(x,y,f)
hold on
contour(X,Y,f,25)
[u,v] = gradient(f);
quiver(x,y,5*u,5*v);

sol = A\b
plot(A\b,'r*');
x0 = [0;0];
disp(['iteracion 0'])
(norm(A*x0-b))^2
plot(x0(1),x0(2),'k*')

colors = {'bs','gs','bo','go','bd','gd'};

syms a fz x;
aux = (norm(A*x-b))^2
fz= matlabFunction(aux)

for i = 1 : 6
    display(['iteracion: ',num2str(i)]);
    (norm(A*x0-b))^2
    alpha = (b-A*x0)'*(b-A*x0)/((b-A*x0)'*A*(b-A*x0));
    x0 = x0 + alpha*(b-A*x0);
    plot(x0,'d')
end



hold off;
```

De esta manera, $x0$ se aproxima a la solución del sistema, esto lo podemos evidenciar con el comportamiento que entrega la gráfica:



Tras las seis iteraciones realizadas, nuestra aproximación toma el valor

$$x^{(6)} = \begin{pmatrix} 1{,}9440 \\ -1{,}9440 \end{pmatrix}$$

El cual es bastante cercano al esperado, considerando que la solución al sistema (2) es

$$x = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

Es decir, la distancia de $x^{(6)}$ a la solución exacta del sistema es aproximadamente 0,0792
Además, la magnitud del residuo es

$$\|r^{(6)}\| = \|Ax^{(6)} - b\| = 0{,}2308$$

(3.8) El sistema resuelto antes es de dos ecuaciones para poder graficar a la función f, así como la dirección de descenso en cada paso del método.
Escriba un nuevo programa, muy similar al segundo, pero que resuelva el siguiente sistema de ecuaciones con el método del máximo descenso

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

En este no pueden graficarse ni f ni sus curvas de nivel. Con $x^{(0)} = (1,1,1,1)^T$¿Cuántas iteraciones realizó el método hasta lograr que $\|Ax - b\| \leq 0{,}01$?

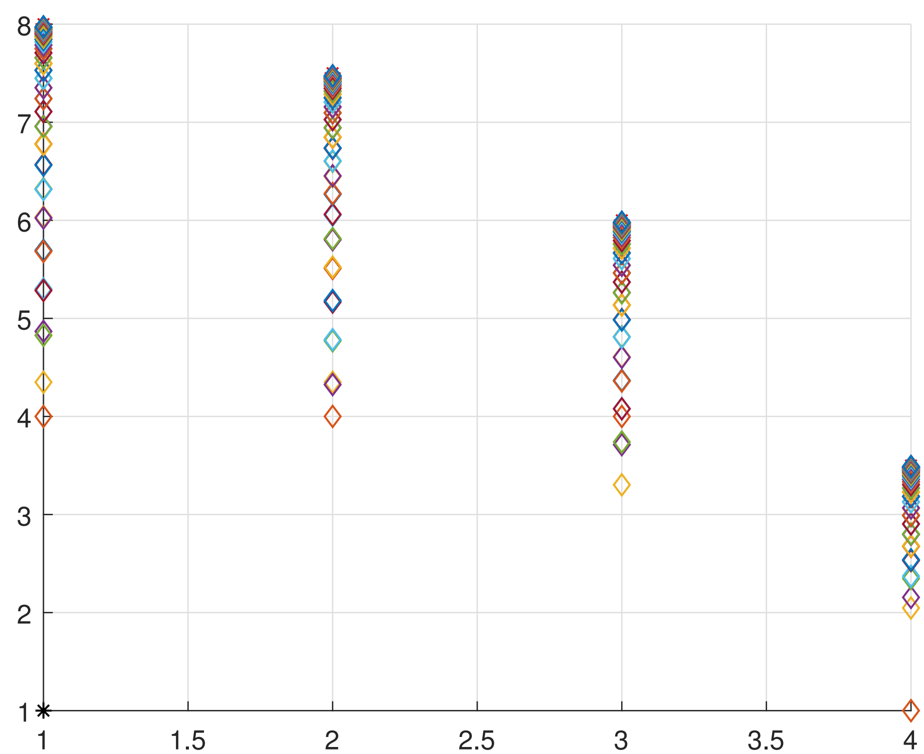
Solución: El programa desarrollado es

Listing 2: Programa 2

```
clc;
% matriz y parte derecha del sistema de ecuaciones lineales
A = [2 -2 0 0;-1 2 -1 0; 0 -1 2 -1; 0 0 -1 2];
b = [1;1;1;1];

N = 20;
x = linspace(-1,5,N);
y = linspace(-5,1,N);
[X,Y] = meshgrid(x,y);
hold on

sol = A\b
plot(A\b,'r*')
x0 = [1;1;1;1];
disp(['iteracion 0'])
(norm(A*x0-b))^2
plot(x0(1),x0(2),x0(3),x0(4),'k*')

colors = {'bs','gs','bo','go','bd','gd'};

```

```
syms a fz x;
aux = (norm(A*x-b))^2;
fz= matlabFunction(aux);

for i = 1:100
    indi = norm(A*x0-b);
    if indi>0.01
        d = -gradient(fz(x0));
        display(['iteracion: ',num2str(i)]);
        disp(indi);
        (norm(A*x0-b))^2
        alpha = (b-A*x0)'*(b-A*x0)/((b-A*x0)'*A*(b-A*x0));
        x0 = x0 + alpha*(b-A*x0);
        plot(x0,'d')
    else
        display(['iteracion: Listo!' num2str(i)]);
        disp(indi);
    end
end



hold off;
```

Si realizamos 100 iteraciones, el programa grafica la siguiente figura



Donde nuestra aproximación, $x^{(100)}$ toma el valor

$$x^{(100)} = \begin{pmatrix} 7{,}9985 \\ 7{,}4987 \\ 5{,}9989 \\ 3{,}4995 \end{pmatrix}$$

El cual es un valor cercano, teniendo en cuenta que la solución al sistema es

$$x = \begin{pmatrix} 8{,}0000 \\ 7{,}5000 \\ 6{,}0000 \\ 3{,}5000 \end{pmatrix}$$

Luego, $\|Ax - b\|$ (residuo del sistema) es menor o igual a 0,01 a partir de la iteración número 64, puesto que

$$\|Ax^{(63)} - b\| = 0{,}0109 \qquad \wedge \qquad \|Ax^{(64)} - b\| = 0{,}0088$$

Problema 4. — Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ tal que $f(x, y, z) = \sin x + 2y + e^{yz}$.[1]

1. Escriba el gradiente y la matriz Hessiana de f en (x, y, z).

   Solución:

$$\nabla f = (\cos x \quad 2 + ze^{yz} \quad ye^{yz})$$

$$H_f = \begin{pmatrix} -\sin x & 0 & 0 \\ 0 & z^2 e^{yz} & (1 + yz)e^{yz} \\ 0 & (1 + yz)e^{yz} & y^2 e^{yz} \end{pmatrix}$$

2. Determine un vector normal a la superficie definida por el gráfico de f en (0,0,0).

   Solución. Sabemos que, al evaluar la expresión del gradiente en algún punto, y éste posee un valor no nulo, entonces el vector resultante es un vector ortogonal a la superficie (Demostrado en ejercicio (3.6)). Así

$$\nabla f(0, 0, 0) = (1, 2, 0) \neq \theta$$

3. Escriba las ecuaciones de la aproximación afín a f en puntos cercanos a (0,0,0) y los polinomios de Taylor de primer y segundo orden de f alrededor de (0,0,0).

   Solución. Aplicando las respectivas definiciones, se tiene que

   - Aproximación afín: (Alrededor del origen)

$$f(x) \approx f(0, 0, 0) + df\big((0, 0, 0); (x - 0, y - 0, z - 0)\big) = 1 + \big\langle \nabla f(0, 0, 0); (x, y, z)\big\rangle = 1 + x + 2y$$

   - Polinomio de Taylor de primer orden:

$$\mathrm{T}_1 := f(0, 0, 0) + (\nabla f \cdot h)(0, 0, 0) = 1 + x + 2y$$

[1] Notamos que f es una función de clase C^∞, por ser una adición de funciones de clase C^∞, por lo que es diferenciable en todo su dominio.

- Polinomio de Taylor de segundo orden:

$$\mathrm{T}_2 := \mathrm{T}_1 + \frac{1}{2!}(\nabla f \cdot h)^2(0,0,0) = 1 + x + 2y + \frac{1}{2}\Big(yz + yz\Big) = 1 + x + 2y + yz$$

4. Utilice los polinomios anteriores para determinar una aproximación al valor de f en $(\frac{1}{10}, \frac{2}{10}, \frac{2}{10})$. Estime el error cometido en cada uno de las aproximaciones.

   Solución. Veamos la aproximación al resultado por cada "método":

   - Aproximación afín:

   $$f\Big(\frac{1}{10}, \frac{2}{10}, \frac{2}{10}\Big) \approx 1 + \frac{1}{10} + \frac{4}{10} = \frac{3}{2} = 1{,}5$$

     Error estimado: Sólo se puede determinar su error restando el resultado real con el resultado aproximado.

   - Polinomio de Taylor de primer orden:

   $$\begin{aligned}\mathrm{T}_1\Big(\frac{1}{10}, \frac{2}{10}, \frac{2}{10}\Big) &= 1 + \frac{1}{10} + \frac{4}{10} + R_2\Big((x,y,z), (\frac{1}{10}, \frac{2}{10}, \frac{2}{10})\Big) \\ &= 1{,}5 + \frac{1}{2!}(2yz)\Big|_{\xi(\frac{1}{10}, \frac{2}{10}, \frac{2}{10})} \\ &= 1{,}5 + \frac{4}{100}\xi^2\end{aligned}$$

     Como $0 < \xi < 1$, entonces $R_2 = \dfrac{4}{100}\xi^2 < 0{,}04$. Así, $\mathrm{T}_1\Big(\frac{1}{10}, \frac{2}{10}, \frac{2}{10}\Big) < 1{,}54$

   - Polinomio de Taylor de segundo orden:

   $$\begin{aligned}\mathrm{T}_2\Big(\frac{1}{10}, \frac{2}{10}, \frac{2}{10}\Big) &= 1 + \frac{1}{10} + \frac{4}{10} + \frac{4}{100} + R_3\Big((x,y,z), (\frac{1}{10}, \frac{2}{10}, \frac{2}{10})\Big) = 1{,}54 + \frac{1}{3!}\big(-x^3\big)\Big|_{\xi(\frac{1}{10}, \frac{2}{10}, \frac{2}{10})} \\ &= 1{,}54 + \frac{1}{6}\Big(-\frac{1}{1000}\xi^3\Big) = 1{,}54 - \frac{1}{6000}\xi^3 = 1{,}54 - 0{,}0002\xi^3\end{aligned}$$

     Con $0 < \xi < 1$, tenemos que $R_3 = -0{,}0002\xi^3 > 0{,}0002$.
     Así, $\mathrm{T}_2\Big(\frac{1}{10}, \frac{2}{10}, \frac{2}{10}\Big) > 1{,}54 - 0{,}0002 = 1{,}5398$

Problema 5. — Dos objetos de masa m están unidos por un resorte. La función $X : \mathbb{R} \longrightarrow \mathbb{R}^3$ describe la posición del primero de ellos en un tiempo y y la función $Y : \mathbb{R} \longrightarrow \mathbb{R}^3$ describe la posición del segundo. Entre ambas se cumple la siguiente relación

$$mX''(t) = k(Y - X) \qquad\qquad mY''(t) = k(X - Y),$$

donde $k \in \mathbb{R}$ es un valor constante que describe la fortaleza del resorte.
Sea $\rho : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ la función que al par de vectores $X, Y \in \mathbb{R}^3$ hace corresponder

$$\rho(X, Y) = \frac{1}{2}k\|X - Y\|^2.$$

1. Suponga que $W : \mathbb{R} \longrightarrow \mathbb{R}^6$ es tal que $\forall t \in \mathbb{R} : W(t) = (X(t), Y(t))$. Demuestre que

   $$mW''(t) = -\nabla\rho(W(t))$$

   Solución. Antes de calcular, notemos lo siguiente:

- Asumiremos la igualdad de aplicaciones $\|\cdot\| = \|\cdot\|_2$ (Norma usual en $\mathbb{R}^n$). De esta forma,

$$\|X-Y\|^2 = \|X-Y\|_2^2 = \left(\sqrt{\sum_{i=1}^{3}(x_i-y_i)^2}\right)^2 = \sum_{i=1}^{3}(x_i-y_i)^2$$

  facilitando así el análisis
- $(X,Y) = \Big((x_1,x_2,x_3),(y_1,y_2,y_3)\Big)$
- $W''(t) = (X''(t), Y''(t))$

Con esto:

$$\begin{aligned}
-\nabla\rho(W(t)) &= -\frac{1}{2}k\Big(2(x_1-y_1),2(x_2-y_2),2(x_3-y_3),-2(x_1-y_1),-2(x_2-y_2),-2(x_3-y_3)\Big)\\
&= -k\Big((x_1,x_2,x_3)-(y_1,y_2,y_3),(y_1,y_2,y_3)-(x_1,x_2,x_3)\Big)\\
&= -k\Big(X-Y,Y-X\Big)\\
&= k\Big(Y-X,X-Y\Big) && \text{(Usando la relación entre } X \text{ e } Y)\\
&= k\Big(\frac{m}{k}X''(t),\frac{m}{k}Y''(t)\Big)\\
&= m\Big(X''(t),Y(t)\Big) = mW''(t)
\end{aligned}$$

2. Demuestre que la cantidad $\frac{1}{2}m\|W'(t)\|^2 + \rho(W(t))$ no cambia con t.

Solución. Una manera de ver que la cantidad dada no varía con respecto al tiempo, es mostrar que su taza de cambio no se ve afectada por el tiempo. Es decir, que si derivamos la expresión con respecto a t, ésta debe dar cero. En efecto

$$\begin{aligned}
\frac{d}{dt}\left(\frac{1}{2}m\|W'(t)\|_2^2 + \rho(W(t))\right) &= \frac{1}{2}m\cdot\frac{d}{dt}\Big(\sum_{i=1}^{3}(x_i')^2 + \sum_{j=1}^{3}(y_j')^2\Big) + \big\langle\nabla\rho(W(t));W'(t)\big\rangle_{\mathbb{R}^3}\\
&= \frac{1}{2}m\Big(2\langle X';X''\rangle_{\mathbb{R}^3} + 2\langle Y';Y''\rangle_{\mathbb{R}^3}\Big) + \big\langle\nabla\rho(W(t));W'(t)\big\rangle_{\mathbb{R}^3}\\
&= m\langle W';W''\rangle_{\mathbb{R}^3} + \langle\nabla\rho(W(t));W'(t)\rangle_{\mathbb{R}^3} && (\star)\\
&= \langle mW''(t);W(t)\rangle_{\mathbb{R}^3} + \langle -mW''(t);W'(t)\rangle_{\mathbb{R}^3} && \text{(Por ítem anterior)}\\
&= 0
\end{aligned}$$

(El paso $(\star)$ es válido pues $\langle W';W''\rangle_{\mathbb{R}^3} = \langle X';X''\rangle_{\mathbb{R}^3} + \langle Y';Y''\rangle_{\mathbb{R}^3}$)