

Optimización II

Tarea III

Joel Chacón Castillo
Guanajuato, México

1. Teoría

Programación lineal es utilizado para modelar y resolver problemas de optimización. Un problema de optimización lineal es definido como se indica a continuación:

$$\min_x f(x) = c^T x, \quad s.a. Ax \geq b, x \geq 0 \quad (1)$$

donde $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$ son vectores constantes. Un aspecto fundamental en la resolución de este tipo de problemas es la complejidad u costo computacional que particularmente están afectados por el número de restricciones y variables. Una categoría de algoritmos para resolver problemas de programación lineal son los métodos de punto interior, que a diferencia al método simplex permiten recorrer el espacio de factible por su interior. Por otra parte se aclara que el método simplex realiza el recorrido por medio de los vértices del simplejo.

1.1. Métodos de punto interior

Considerando el siguiente problema de programación lineal en la forma estándar:

$$\min_x c^T x, \quad s.a. Ax \geq b, x \geq 0 \quad (2)$$

donde $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$ son vectores constantes. Además A es una matriz de rango completo por columnas. También se parte del hecho de que A cumple las condiciones LICQ. La forma dual de este problema es como se indica a continuación:

$$\max_{s, \lambda} b^T \lambda, \quad s.a. A^T \lambda + s = c, s \geq 0 \quad (3)$$

donde $\lambda \in \mathbb{R}^m$ y $s \in \mathbb{R}^n$ son vectores que corresponden a las variables duales.

Además la condiciones de Karusck-Kuhn-Tucker (KKT) se definen a continuación:

$$A^T \lambda = c \quad (4)$$

$$Ax \lambda = b \quad (5)$$

$$s^T \lambda = 0 \quad (6)$$

$$x, s \succeq 0 \quad (7)$$

En particular los métodos primal-dual son diseñados para encontrar las soluciones x^*, λ^*, s^* del problema de programación lineal. Específicamente, estos procedimientos utilizan un mecanismo de optimización similar a los métodos de Newton de la forma $x^{t+1} = x^t + \alpha p$ donde x^t es el vector a optimizar, α es considerado como el tamaño de paso y p es la dirección de descenso. En general las condiciones de optimalidad se pueden resolver de forma sencilla en base a su forma lineal, sin embargo existe un inconveniente con la condición de no negatividad $x, s \succeq 0$, pues esta condición representa un sistema no lineal y por lo tanto una dificultad importante en los métodos de punto interior.

En la derivación de los métodos de punto interior primal-dual y en base a las condiciones KKT se propone un funcional que tiene el efecto de realizar un mapeo $F : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$, esto se realiza como se indica a continuación:

$$F(x, \lambda, s) = \begin{pmatrix} A^T \lambda + s - x \\ Ax - b \\ XSe \end{pmatrix} = \mathbf{0} \quad (8)$$

donde $X = \text{diag}(x_1, x_2, \dots, x_n)$, $S = \text{diag}(s_1, s_2, \dots, s_n)$ y $e = \mathbf{1} = (1, 1, \dots, 1)^T$. El paso α se determina a partir de la función $F(x, \lambda, s)$. Además se utiliza el método de Newton en el cual la dirección de descenso se satisface en:

$$\nabla^2 F(x, \lambda, s) \Delta_{x, \lambda, s} x = -\nabla F(x, \lambda, s) \quad (9)$$

Los métodos de tipo primal-dual realizan iteraciones en x, λ, s de tal forma que satisfacen las condiciones de desigualdad de forma estricta. Por lo tanto para resolver la ecuación $F(x, \lambda, s) = 0$ por medio del método de Newton se requiere determinar el tamaño de paso y tener una medida de ajusta el cual indica que tan apropiado es un vector en el espacio de búsqueda. Una

medida de factibilidad consiste en el promedio $\mu = \frac{x^T s}{n} = \frac{1}{n} \sum_{i=1}^n x_i s_i$ mejor conocida como la medida de dualidad.

Algorithm 1 Primal-Dual

1: Dado (x^0, λ^0, s^0) con $(x^0, s^0) > 0$
2: **for** $k=1,2,\dots$, **do**
3: Resolver

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta_x \\ \Delta_\lambda \\ \Delta_s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe + \sigma\mu e \end{bmatrix} \quad (10)$$

4: donde $\sigma \in [0, 1]$, $\mu = x^T s / n$, $r_c = A^T \lambda + s - c$, $r_b = Ax - b$
5: $(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k, \lambda^k, s^k) + \alpha(\Delta_x, \Delta_\lambda, \Delta_s)$
5: Seleccionar α de modo que $(x^{k+1}, s^{k+1}) > 0$

Es importante considerar que éste método únicamente requiere como puntos iniciales $(x^0, s^0) > 0$, para realizar esto se puede aplicar varias heurísticas. Además es importante mencionar que la elección del parámetro de centrado σ y la longitud de paso α son muy importantes para el rendimiento del algoritmo.

1.1.1. Ruta central

El conjunto factible primal dual F y el conjunto estrictamente factible F^0 se definen como se indica a continuación:

$$F = \{(x, \lambda, s) | Ax = b, A^T \lambda + s = c, (x, s) \geq 0\} \quad (11)$$

$$F^0 = \{(x, \lambda, s) | Ax = b, A^T \lambda + s = c, (x, s) > 0\} \quad (12)$$

La ruta central C es un arco de puntos factibles estrictos de gran importancia en los algoritmo primal-dual.

$$\begin{aligned} A^T \lambda &= c \\ Ax &= b \\ x_i s_i &= \tau \\ (x, s) &> 0 \end{aligned} \quad (13)$$

Por lo tanto se define a la ruta central como $C = \{(x_\tau, \lambda_\tau, s_\tau) | \tau > 0\}$ Además al considerar τ como parte de las KKT el problema es modificado un problema de penalización, particularmente este es el caso del problema de barrera logarítmica $\min c^T x - \sum_{i=1}^n \ln(x_i) \quad s.a. \quad Ax = b$

1.2. Long-Step Path Following

En este método se considera el vecindario $N_\infty(\gamma)$ que es definido de la siguiente forma:

$$N_\infty(\gamma) = \{(x, \lambda, s) \in F^0 | x_i s_i \geq \gamma \mu \forall i = 1, 2, \dots, n\} \quad (14)$$

Los métodos de path-following reducen todos los productos $x_i s_i$ a cero en una razón menor o igual que el algoritmo primal-dual, esto se realiza manteniendo todas las iteraciones en esta vecindad. Además este algoritmo realiza un progreso rápido por el uso del vecindario, para un γ cercano a cero. Este algoritmo también depende de los dos parámetros σ_{min} y σ_{max} también considerados como cotas inferior y superior en el parámetro de centrado σ . Al dirección de búsqueda que se toma es la misma que la obtenida en el algoritmo Primal-Dual, y el tamaño de paso es tan grande como sea permitido por el vecindario $N_\infty(\gamma)$.

Un inconveniente principal de este método es que requiere de un punto inicial dentro de la vecindad y esto puede ser computacionalmente costoso, pues se requiere que se cumplan las condiciones KKT. Una forma para encontrar el punto inicial (que se utiliza en este trabajo) consiste en primero ejecutar el método Primal-Dual únicamente minimizando r_c, r_b con el propósito de generar un punto que pertenezca al vecindario $N_\infty(\gamma)$, posteriormente se utilizan los puntos generados (x, λ, s) en el método de Long-Step-Path-Following.

Algorithm 2 Long-Step-Path-Following

- 1: Dado $\gamma, \sigma_{min}, \sigma_{max}$ con $gamma \in (0, 1)$ y $0 < \sigma_{min} \leq \sigma_{max} < 1$ y (x^0, λ^0, s^0) con $(x^0, s^0) \in N_\infty(\gamma)$
 - 2: **for** $k=1, 2, \dots$, **do**
 - 3: Elegir $\sigma \in [\sigma_{min}, \sigma_{max}]$
 - 4: Resolver

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta_x \\ \Delta_\lambda \\ \Delta_s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe + \sigma \mu e \end{bmatrix} \quad (15)$$
 - 5: donde $\sigma \in [0, 1]$, $\mu = x^T s / n$, $r_c = A^T \lambda + s - c$, $r_b = Ax - b$
 $(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k, \lambda^k, s^k) + \alpha(\Delta_x, \Delta_\lambda, \Delta_s)$
 - 6: Seleccionar α como el valor más largo en $[0, 1]$ de modo que $(x^{k+1}, \lambda^{k+1}, s^{k+1}) \in N_\infty(\gamma)$
-

1.3. Predictor - Corrector - Mehrotra

El método Predicto-Corrector consiste en un procedimiento similar al Primal-Dual, sin embargo su peculiaridad es ajustar el tamaño de paso tanto para las variables primales como para las variables duales, como ventaja este

algoritmo requiere menos parámetros de usuario con el precio de resolver dos sistemas de ecuaciones en lugar de uno.

Algorithm 3 Predicto-Corrector Mehrotra

- 1: Dado $(x^0, s^0) > 0$
2: **for** $k=1,2,\dots$, **do**
3: Resolver

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{aff} \\ \Delta \lambda^{aff} \\ \Delta s^{aff} \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe \end{bmatrix} \quad (16)$$

- donde $r_c = A^T \lambda + s - c$, $r_b = Ax - b$
4: $(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k, \lambda^k, s^k) + \alpha(\Delta x, \Delta \lambda, \Delta s)$
5: Calcular α_{aff}^{pri} , α_{aff}^{dual} y μ_{aff} de la siguiente forma

$$\begin{aligned} \alpha_{aff}^{pri} &= \min(1, \min_{i: \Delta x_i^{aff} < 0} -\frac{x_i}{\Delta x_i^{aff}}) \\ \alpha_{aff}^{dual} &= \min(1, \min_{i: \Delta s_i^{aff} < 0} -\frac{s_i}{\Delta s_i^{aff}}) \\ \mu_{aff} &= (x + aff^{pri} \Delta x^{aff})^T (s + aff^{dual} \Delta s^{aff}) / n \\ \sigma &= (\frac{\mu_{aff}}{\mu})^3 \end{aligned} \quad (17)$$

- 6: Resolver

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^{aff} \\ \Delta \lambda^{aff} \\ \Delta s^{aff} \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe - \Delta X^{aff} \Delta S^{aff} e + \sigma \mu e \end{bmatrix} \quad (18)$$

- donde $r_c = A^T \lambda + s - c$, $r_b = Ax - b$
7: Calcular $\alpha_{max}^{pri} = \min_{i: \Delta x_i < 0} -\frac{x_i}{\Delta x_i}$ y $\alpha_{max}^{dual} = \min_{i: \Delta s_i < 0} -\frac{s_i}{\Delta s_i}$
8: Asignar $\alpha^{pri} = \min(1, \eta \alpha_{max}^{pri})$ y $\alpha^{dual} = \min(1, \eta \alpha_{max}^{dual})$ donde $\eta \in [0, 9, 1, 0]$
9: Actualizar $x = x + \alpha^{pri} \Delta x$ y $(\lambda, s) = (\lambda, s) + \alpha^{dual} (\Delta \lambda, \Delta s)$
-

Es importante notar que este método únicamente requiere vectores $(x, s) \succeq 0$. Además, en general el número de iteraciones necesarias deberá ser considerablemente menor al método Primal-Dual y Long-Step-Path-Following, ya que se realiza el cálculo del tamaño de paso más adecuado. Una desventaja importante es que a diferencia de los métodos Primal-Dual y Long-Step-Path-Following, no existe una demostración teórica de convergencia, mas bien en la literatura se han encontrado casos en que el algoritmo diverge [1]. Una ventaja es el hecho de que no es necesario proporcionar muchos parámetros de entrada $(x^0, s^0) > 0$ y ρ .

1.4. Métodos para escoger un punto inicial

El desempeño de los métodos de punto-interior están afectados por el punto inicial, inclusive un método puede no llegar a converger a una solución si el punto inicial no cumple con las condiciones KKT, como es el caso del método Long-Step-Path-Following. A continuación se describen tres métodos para realizar encontrar un punto inicial adecuado.

1.5. Método de inicialización de dos fases

En este método primero se resuelve un problema más sencillo de la forma

$$\begin{aligned} \min_{x,u} \quad & u \\ \text{s.a.} \quad & \\ Ax + (b - Ax^0)u = b \\ & x, u \succeq 0 \end{aligned} \tag{19}$$

Posteriormente se puede obtener $\lambda = (AA^T)^{-1}Ac$, $s = c - A^T$, es importante notar que el vector s puede tener entradas negativas. o de forma alternativa se resuelve su correspondiente forma dual

$$\begin{aligned} \min_{\lambda,s,u} \quad & -u \\ \text{s.a.} \quad & \\ A^T\lambda + s + (c - A^T\lambda^0 + s^0)u = c \\ & \lambda, s, u \succeq 0 \end{aligned} \tag{20}$$

1.6. Método de inicialización equivalente al problema de norma mínima

Este tipo de problema consiste en resolver de forma teórica los siguiente problemas

$$\min_x \frac{1}{2} \|x\|^2 \quad \text{s.a.} Ax = b \tag{21}$$

$$\min_{\lambda,s} \frac{1}{2} \|s\|^2 \quad \text{s.a.} A^T\lambda + s = c \tag{22}$$

La solución de los problemas anteriores corresponde a

$$\begin{aligned} \hat{x} &= A^T(AA^T)^{-1}b \\ \hat{\lambda} &= (AA^T)^{-1}Ac \\ \hat{s} &= c - A^T\hat{\lambda} \end{aligned} \tag{23}$$

Debido a que algunas de las componentes de (\hat{x}, \hat{s}) pueden ser nevativos, para garantizar que sean no negativas se definen como

$$\begin{aligned}\delta_x &= \max(0, -(3/2)\min_i \hat{x}_i) \\ \delta_s &= \max(0, -(3/2)\min_i \hat{s}_i)\end{aligned}\tag{24}$$

Posteriormente, el punto inicial se calcula de la siguiente forma:

$$\begin{aligned}x^0 &= \hat{x} + \delta_x e \\ s^0 &= \hat{s} + \delta_s e \\ \lambda^0 &= \hat{\lambda}\end{aligned}\tag{25}$$

donde $\hat{\delta}_x, \hat{\delta}_s$ son los promedios pesados

$$\begin{aligned}\hat{\delta}_x &= \frac{1}{2} \frac{\hat{x}^T \hat{s}}{\mathbf{1}^T \hat{s}} \\ \hat{\delta}_s &= \frac{1}{2} \frac{\hat{x}^T \hat{s}}{\mathbf{1}^T \hat{x}}\end{aligned}\tag{26}$$

esto último es para asegurar que x^0, s^0 no sean muy distintos y además no sean cercanos a cero.

1.7. Método de inicialización por medio del espacio nulo - ideada por mi

En este procedimiento se asume que el valor x del espacio factible se puede obtener al resolver $Ax = b$, sin embargo dado que el sistema es $m \times n$ donde $n \geq m$ se puede obtener distintos x que cumplan $Ax = b$.

Por lo tanto dado que la matriz A es de rango completo por columnas, se pueden seleccionar m columnas de A linealmente independientes mejor indicado como A^* , de esta forma se puede obtener la solución x^* al sistema $A^*x^* = b$ que también será solución de $Ax^* = b$ (nota A es $m \times n$ y A^* es $m \times m$). Posteriormente se obtiene una solución de sistema homogéneo es decir obtener un x^{null} de $Ax^{null} = 0$ donde $x^{null} \neq 0$.

Finalmente se observa que si se resuelve $Ax^* = b$ y x^* tiene entradas negativas se le puede sumar un múltiplo de la solución que corresponde al conjunto homogéneo. Es decir $x^0 = x^* + tx^{null}$ donde las entradas de x^0 son positivas, esto es válido ya que $A[x^* + tx^{null}] = b = Ax^* + A(tx^{null}) = Ax^* = b$. Por lo tanto el problema consiste en encontrar un t tal que todas las entradas de $x^* + tx^{null}$ sean positivas. Curiosamente esta forma es similar a la forma de Newton donde t podría ser el tamaño de paso y x^{null} la dirección de descenso. Ahora el problema es encontrar un valor de t que haga $x^* + tx^{null} \succ 0$, esto se logra con $t = \min_{i: x_i^{null} < 0} -\frac{x_i^*}{x_i^{null}}$.

2. Análisis experimental

En esta validación experimental se generan y resuelven de forma aleatoria problemas de programación lineal en su forma estándar para $n = 100 * i$ con $i \in \{1, 2, \dots, 10\}$ y $m = 80$, con 30 ejecuciones con cada configuración. Los parámetros que se utilizaron se especifican a continuación:

- **Primal-Dual:** $\alpha = 0,5$.
- **Long-Step-Path-Following:** $\sigma_{min} = 0,001$, $\sigma_{max} = 0,01$ y $\gamma = 1e - 3$.
- **Mehrotra:** $\eta = 0,95$

El criterio de paro en todos los algoritmos de punto interior fue con un límite de 20 iteraciones y $||\Delta x|| < 1e - 5$. En la figura 1 se presenta la media del número de iteraciones, media del valor de la función objetivo y media del tiempo requerido por cada algoritmo. Se puede observar que el método del Simplex es el que involucra muchas iteraciones, sin embargo cada iteración requiere un tiempo computacional menor que los métodos de punto interior. Todos los algoritmos se fueron aplicados al mismo problema de cada ejecución, principalmente se puede observar que el valor de la función objetivo para $n = 100$ es mejor por los algoritmos de punto interior que por el método simplex. Además se observa que al considerar 1000 variables los métodos Long-Step-Path-Following y Mehrotra empeoran de forma significativa, posiblemente esto se debe a que el número de iteraciones (20 en este caso) no fue suficiente para proporcionar un valor de la función objetivo suficientemente aceptable.

Como es de esperarse el método de Mehrotra requiere tiempo computacional elevado. El método Long-Step-Path-Following requiere el mayor tiempo computacional ya que el método para generar el punto inicial $x^0 \in N_{\infty}(\gamma)$ requiere de un tiempo elevado. Además se puede observar que en algunas situaciones el método Primal-Dual requiere un menor tiempo computacional que el Simplex y además proporciona un mejor valor de la función objetivo, esto se puede rectificar en las tablas 1.

En la figuras 2 se presenta la desviación estándar de los 30 resultados de cada configuración, principalmente se puede observar que en el caso del Simplex la desviación estándar del número de iteraciones es muy distinto, esto se debe a que el Simplex puede requerir de muchas iteraciones para llegar al punto óptimo, ya que su mecanismo consiste en seleccionar aristas del polítopo conformado por las restricciones del problema. Finalmente, se

observa que el tiempo computacional requerido por el algoritmo Long-Step-Path-Following tiende a ser muy distinto, esto se debe al procedimiento para encontrar el punto inicial que pertenezca al vecindario.

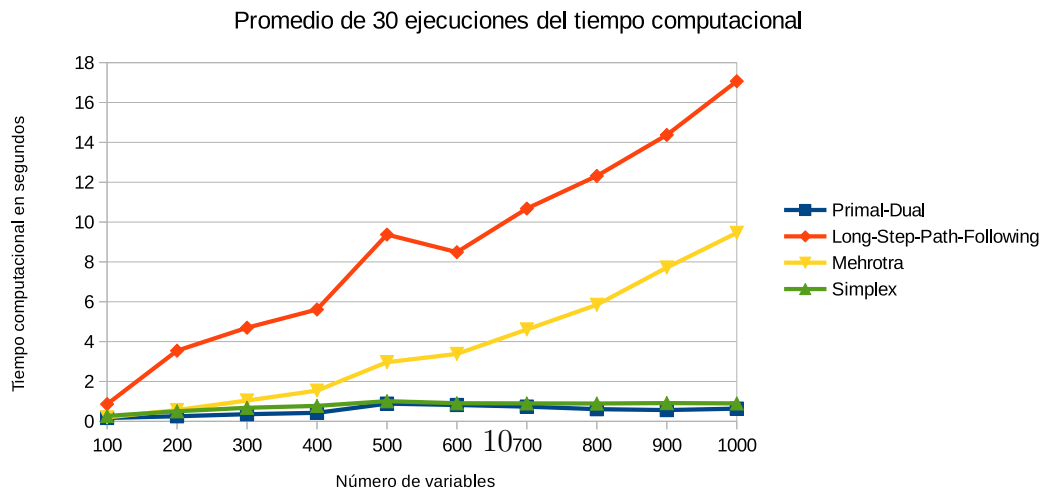
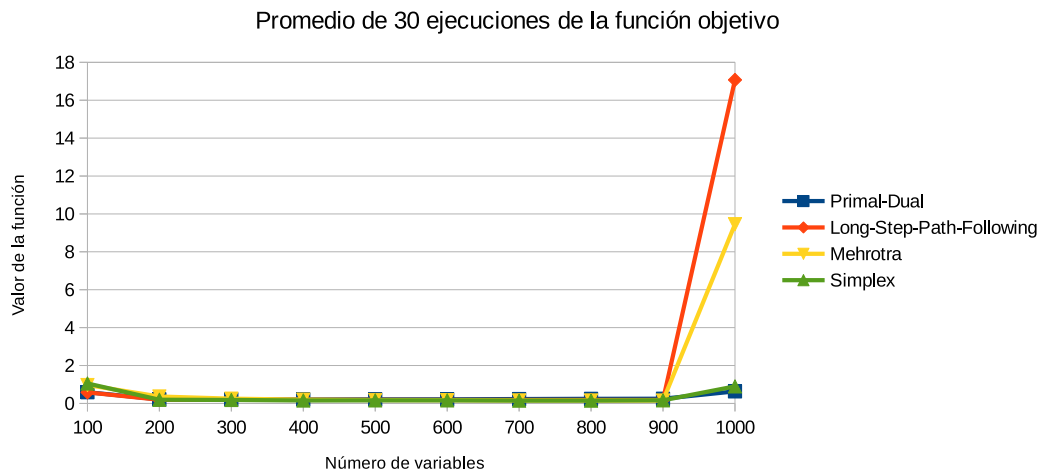
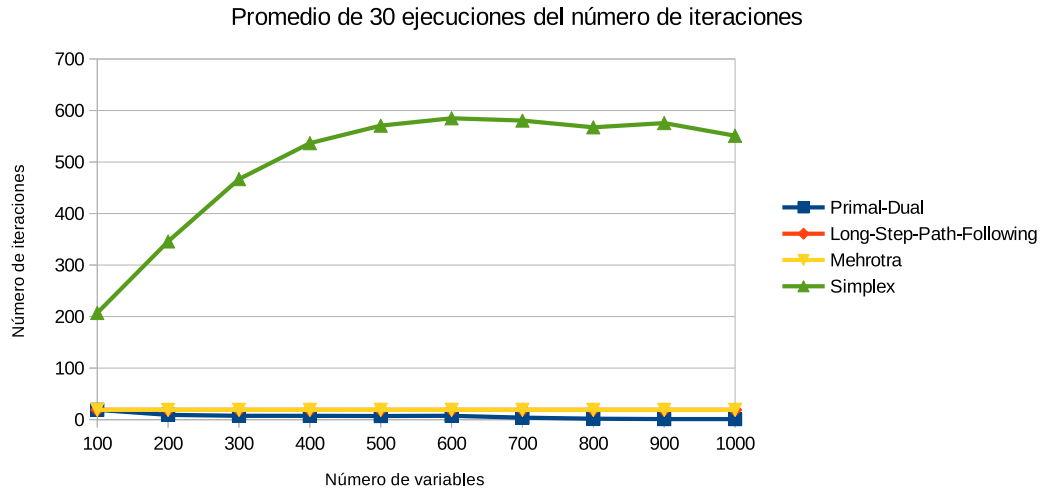


Figura 1: El número de restricciones consideradas fueron de $m = 80$.

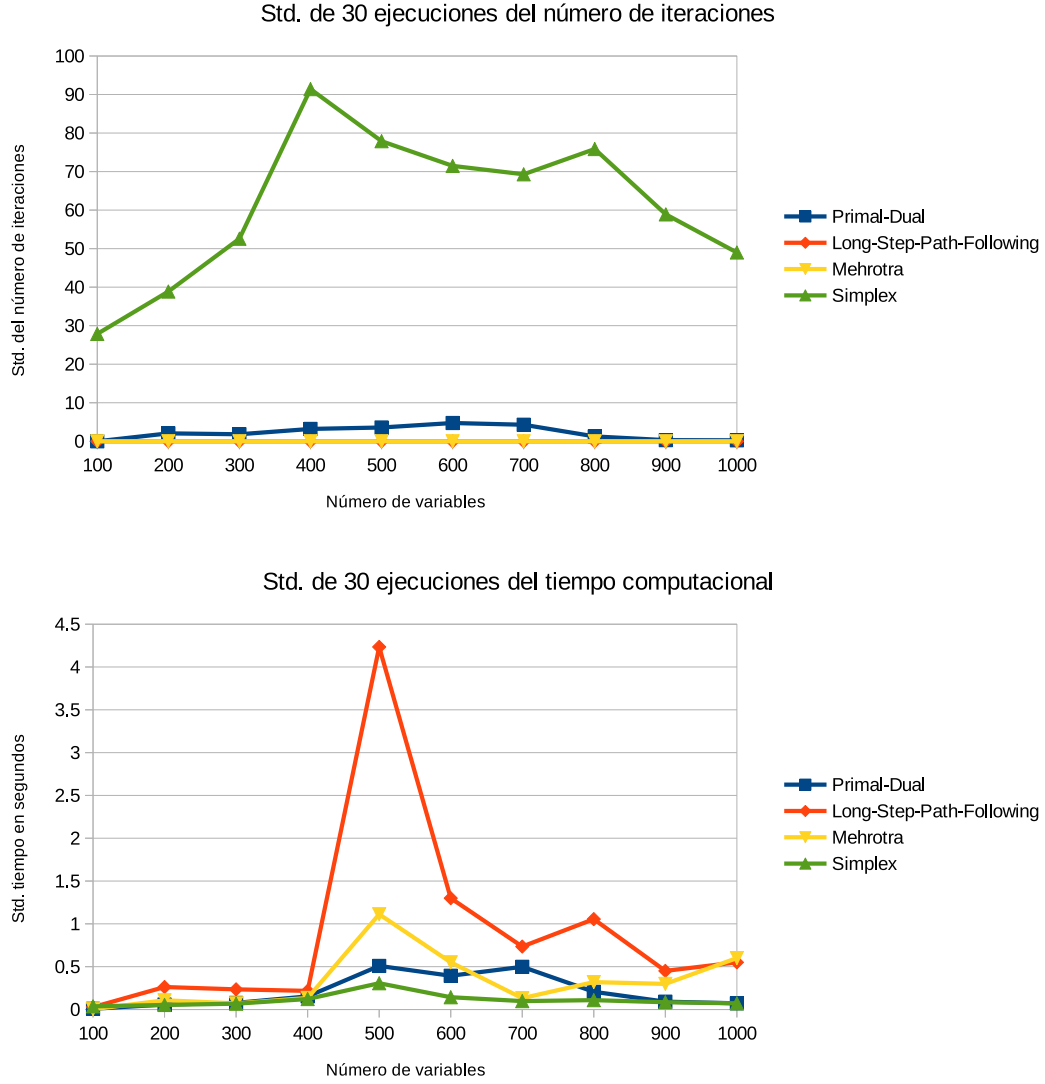


Figura 2: El número de restricciones consideradas fueron de $m = 80$.

3. Análisis de parámetros de centrado en los métodos de punto interior

En la figura 3 se presentan los valores normalizados que corresponden al parámetro de centrado en los algoritmos de punto interior Primal-Dual, Long-

Cuadro 1: Média del número de iteraciones, valor de función objetivo y tiempo de las 30 ejecuciones realizadas

<i>Media de iteraciones</i>										
<i>N</i>	100	200	300	400	500	600	700	800	900	1000
<i>Primal-Dual</i>	19	9.3	7.433333333	7.366666667	7.033333333	7.4	3.866666667	1.766666667	1.1	1.066666667
<i>Long-Step-Path-Following</i>	19	19	19	19	19	19	19	19	19	19
<i>Mehrotra</i>	19	19	19	19	19	19	19	19	19	19
<i>Simplex</i>	207	345.8333333	466.8333333	536.6	570.6	584.8666667	580.4333333	567.3333333	575.5333333	551.0666667
<i>Media de la función objetivo</i>										
<i>N</i>	100	200	300	400	500	600	700	800	900	1000
<i>Primal-Dual</i>	0.5916970002	0.2139867805	0.2153545298	0.2152102171	0.2178102007	0.217243397	0.2320369873	0.2413336018	0.2466227145	0.6373060385
<i>Long-Step-Path-Following</i>	0.583763836	0.1930506566	0.1889859624	0.186413086	0.1850146869	0.1841476891	0.1831888034	0.182659612	0.182111192	17.0697051446
<i>Mehrotra</i>	0.9685721969	0.3638375414	0.2499114915	0.2001972765	0.181107844	0.1734810266	0.1587911806	0.1558558198	0.1513754161	9.4576407512
<i>Simplex</i>	1.0515155543	0.1930504912	0.1826714586	0.155442998	0.1607028168	0.1608525768	0.149228438	0.1485505237	0.1625557258	0.8983331601
<i>Media del tiempo computacional</i>										
<i>N</i>	100	200	300	400	500	600	700	800	900	1000
<i>Primal-Dual</i>	0.1698413769	0.2536817551	0.3581613938	0.4276407639	0.8870601813	0.8197888772	0.7370838165	0.6109750827	0.5647231658	0.6373060385
<i>Long-Step-Path-Following</i>	0.8632011573	3.5436320067	4.6972787221	5.6114912748	9.3732870817	8.4874481996	10.6776926994	12.3141242027	14.3703053077	17.0697051446
<i>Mehrotra</i>	0.1979612112	0.568127044	1.0436141332	1.5385317087	2.9670418104	3.3811044057	4.6091835976	5.8444034815	7.7148765405	9.4576407512
<i>Simplex</i>	0.2699072361	0.5030421575	0.678675429	0.7756952604	1.00984629	0.9074576139	0.8988627831	0.8934083939	0.9162847519	0.8983331601

Step-Path-Following y Mehrotra. Particularmente se observa que el método Primal-Dual decremanta de forma paulatina, por otra parte se observa que el método de Mehrotra decremanta este valor de forma significativa en la segunda iteración y en las iteraciones posteriores no es muy signifcativo el cambio, esto indica que el método de Mehrotra converge de forma acelerada, sin embargo este último método requiere de un tiempo computacional elevado. Finalmente, se puede observar que el método de Long-Step-Path-Following inicia con valores muy pequeños, esto se debe a que la solución inicial que se proporciona debe estar en en un vecindario definido y por lo tanto los valores de centrado se límitan a los definidos por la ruta de centrado que en contraste a los demás métodos son mucho menores.

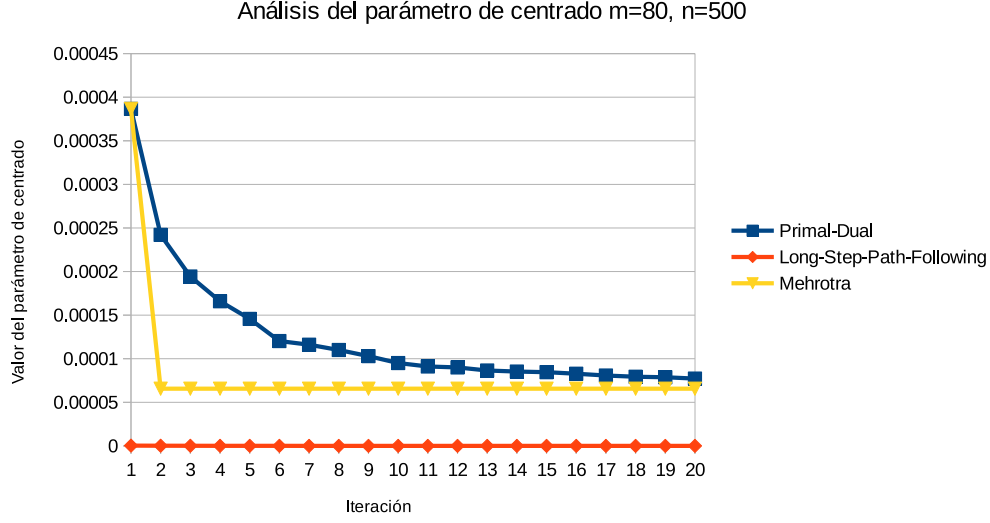


Figura 3: Análisis del parámetro de centrado (μ) con el algoritmo Primal-Dual, Long-Step-Path-Following y Mehrotra.

En la figura 4 se observa el comportamiento de los parámetros de centrado μ y σ , primero se hace mención que éstos resultados están normalizados, que aunque el valor de μ es muy pequeño parece elevado por el efecto de la normalización. Se probaron varias configuraciones y se presentan los resultados más representativos. Es importante aclarar que en el caso anterior se observa como el μ decrementa conforme itera el algoritmo, en este caso ocurre lo contrario debido a que el problema que se considera en este caso es más sencillo y por lo tanto el punto inicial no genera valores elevados, además se hace notar que los valores están normalizados. El propósito de esta figura es mostrar la tendencia de μ con respecto a σ . En particular, se desea mostrar la relación del comportamiento entre μ y σ , recordando que $\mu_{aff} = (x + \alpha_{aff}^{pri} \Delta_x^{aff})^T (s + \alpha_{aff}^{dual} \Delta_s^{aff}) / n$, $\mu = x^T s / n$ y $\sigma = (\frac{\mu_{aff}}{\mu})^3$, se puede observar que la relación que existe entre μ_{aff} y μ tiende a la unidad, posiblemente esto se debe a que el procedimiento convergió a un punto y por lo tanto la medida de factibilidad es constante.

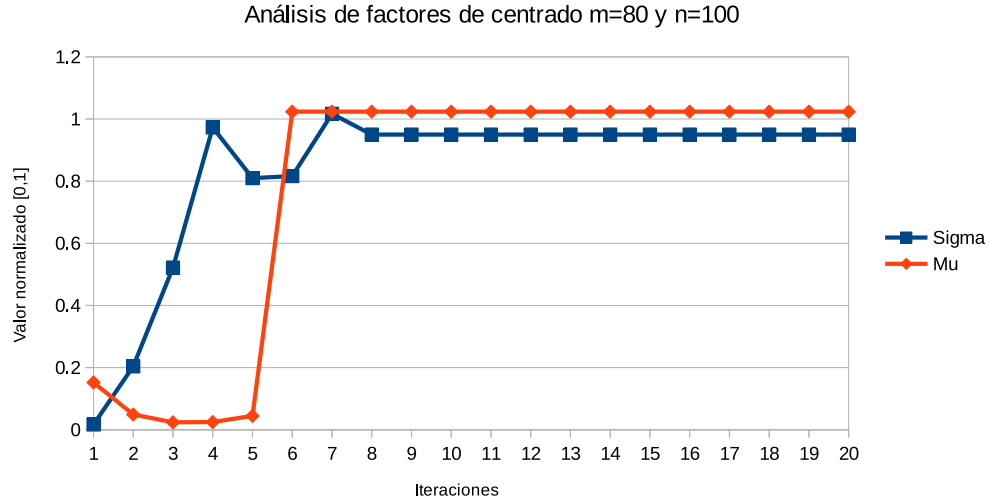


Figura 4: Análisis del parámetro de centrado con el algoritmo Mehrotra (normalizados).

En general se puede decir que μ es considerado como una medida de dualidad y por lo tanto su comportamiento dependerá de las características del problema y del punto inicial. En los problemas difíciles el punto inicial tenderá a producir una medida de dualidad elevado el cual irá decrementándose hasta cierto punto conforme itera el algoritmo, sin embargo si el punto inicial tiene una medida de dualidad baja que es el caso de los problemas sencillos éste puede incrementar en base a el tradeoff de la función objetivo, es decir, esta medida puede empeorar muy poco pero con una mejora significativa de la función objetivo, de hecho esto es lo más importante de los métodos de punto interior ya que se puede explorar la región no factible y el interior del polítopo.

- 1 [1] J. Nocedal, S. J. Wright, Sequential quadratic programming, Springer,
- 2 2006.