

# SAA – SAMPLE AVERAGE APPROXIMATION METHOD, APLICADO A LA SOLUCIÓN DE MODELOS DE PROGRAMACIÓN LINEAL ESTOCÁSTICOS

*Héctor Hernan Toro Díaz*

---

*Profesor Pontificia Universidad Javeriana, Colombia.  
Contacto: hetordi@pino.univalle.edu.co*

## **Resumen**

Este trabajo aborda la solución de problemas de optimización de programación lineal estocásticos, mediante la generación de escenarios aleatorios usando el método monte carlo. Se presenta el algoritmo de solución, así como algunos de sus detalles de implementación computacional. Se desarrolla un ejemplo ilustrativo de la aplicación del algoritmo, sobre un modelo típico de PL Mixta, conocido como ULSP – *Uncapacitated Lot Sizing Problem*. Se presenta los resultados computacionales que permiten concluir, desde una perspectiva práctica, sobre la convergencia del algoritmo hacia la solución óptima del problema de PL.

**Palabras clave:** optimización estocástica, generación aleatoria de escenarios, Método Monte Carlo.

## **Abstract**

This work faces the solution process of stochastic linear programming problems, using random scenario generation provided by Monte Carlo method. The solution algorithm is presented, as well as some of its computational implementation details. An example that shows the applicability of the algorithm is developed, using a typical LP problem, usually referred as ULSP, uncapacitated lot sizing problem. Computational results are shown, and they allow, from a practical perspective, to conclude about algorithm convergence to the optimal solution of LP problem.

**Keywords:** stochastic optimization, random scenario generation, Monte Carlo Method.

## 1. Introducción

En el estudio de los actuales sistemas de producción y distribución, cuando se abordan desde la perspectiva de una cadena de abastecimiento con la intención de tomar buenas decisiones a nivel estratégico, es usual que aparezcan modelos matemáticos de programación lineal mixta como herramientas de soporte a la toma de decisiones. Estos modelos suelen ser de gran tamaño, incluyendo típicamente cientos de miles de variables, y un número similar de restricciones, incluso para cadenas de abastecimiento de carácter regional.

Varios trabajos han abordado la modelación matemática de cadenas de abastecimiento, tanto regionales como internacionales, incluyendo aplicaciones prácticas. Entre ellos pueden citarse Eksioglu (2000), Goetschalckx et. al. (2002), Vidal y Goetschalckx (2001), Toro (2001) y Bravo y Bravo (2005). Elementos comunes de estos trabajos es el abordaje del problema de optimización de la cadena de suministro mediante la modelación matemática; que se ha llegado en todos los casos a modelos de PL mixta, o bien, a modelos no lineales que se pueden solucionar mediante un esquema que los descompone en un conjunto de modelos de PL Mixta relacionados; y por último, que los modelos son de naturaleza determinística, es decir, los parámetros que los asocian a una situación particular se asumen con condición de certeza.

Ha resultado más o menos natural que luego de trabajar con tal esquema de aproximación al problema de optimización de las redes de suministro se proponga estudiar el impacto de elementos de variabilidad en la estructura de los modelos, hecho que ha sucedido al reconocer que varios parámetros críticos que se han modelado de modo determinístico, son en realidad variables aleatorias, de las cuales apenas puede hacerse un pronóstico con un cierto nivel de confianza, respecto a lo que sería su valor o valores futuros. La inclusión de tales parámetros como variables aleatorias, cambia el panorama algorítmico de solución de los modelos matemáticos, incluso si se mantiene toda su restante estructura. En tal sentido, se ha visto la necesidad de estudiar los métodos de solución que pueden ser aplicados sobre modelos de programación matemática que incluyen elementos estocásticos, y hacer estos estudios y desarrollos en paralelo al proceso de plantear modelos de cadenas de abastecimiento que incluyen cada vez más elementos de representación del mundo real.

El presente trabajo aborda el estudio de uno de esos métodos de solución, bastante prometedor tanto por la solidez teórica con la cual se ha desarrollado, que incluye pruebas de optimalidad, análisis de convergencia y respuesta computacional, como por su relativa facilidad de implementación, usando para ello herramientas que usualmente están disponibles en el panorama de toma de decisiones en el mundo real. El método, conocido como SAA, *Sample Average Approximation*, se define básicamente como un esquema de aproximación mediante simulación monte carlo, al problema de optimización de modelos de programación matemática que incluyen parámetros estocásticos en su definición. La idea básica del método es que se genera una muestra aleatoria de los parámetros del modelo, con lo cual resulta una realización particular del modelo cuya naturaleza sería determinística, y para la cual se conocen algoritmos de optimización eficientes. Este proceso se repite un número determinado de veces, y el valor esperado de la función objetivo es aproximado mediante la función promedio aplicada sobre los valores de la función objetivo que se obtiene de cada muestra, luego del proceso de optimización. El número de veces que se repite el proceso es usualmente dinámico, pues se prefiere realizar la ejecución de una nueva iteración hasta que se cumpla un cierto criterio de parada, predeterminado por el decisor.

## 2. Problema de Optimización

La caracterización general de los problemas de optimización que se abordan en este trabajo es la siguiente:

$$\min_{x \in S} \{g(x) = E[G(x, W)]\} \quad (1)$$

En la ecuación (1),  $\mathbf{W}$  es un vector aleatorio que tiene distribución de probabilidad conocida  $\mathbf{P}$ ;  $S$  es un conjunto finito (por ejemplo un subconjunto de  $\mathbb{R}^n$  con coordenadas enteras);  $G(x, w)$  es una función real de dos vectores de variables,  $x$  y  $w$ ; y  $E[G(x, W)] = \int G(x, w)P(dw)$  es el correspondiente valor esperado. Se asume que la función de valor esperado  $g(x)$  está bien definida, de modo que para cada  $x \in S$  la función  $G(x, \bullet)$  se puede evaluar, y que el valor es finito, es decir,  $E\{|G(x, W)|\} < \infty$ .

El problema general es interesante de estudiar cuando además se dan una serie de características particulares que lo convierten en un problema de mayor complejidad, a saber:

1. La función de valor esperado  $g(x) = E[G(x, W)]$  no puede escribirse fácilmente de forma analítica, o bien, incluso en tal caso los valores no se pueden calcular de un modo eficiente.
2. En cambio, la función  $G(x, w)$  es fácilmente evaluable dados valores particulares para  $x$  y  $w$ .
3. El conjunto de soluciones factibles  $S$ , aunque es finito, es muy grande, de modo que cualquier estrategia de enumeración exhaustiva de soluciones queda descartada debido al consumo de recurso computacional. Es usual, de hecho, que en problemas de programación matemática con variables enteras el número de soluciones que conforman la región factible crezca exponencialmente respecto al número de variables del problema.

Un problema de optimización con variables enteras es usualmente más difícil de resolver que su contraparte con variables continuas, lo cual sucede incluso en un panorama determinístico. Para el caso del problema que se ha descrito, la otra dificultad es que la función objetivo  $g(x)$  puede ser muy complicada o difícil de evaluar, incluso si se usan métodos numéricos de aproximación.

Se ha publicado variados trabajos que abordan el problema de optimización que nos concierne, la mayoría de los cuales trabaja con casos en los que el conjunto de soluciones factibles es lo suficientemente pequeño, de modo que se puede hacer una estimación de la función  $g(x)$ , para cada  $x$  dentro de la región factible. Como ejemplo de tales trabajos, Kleywegt et al (2001) cita a Hochberg y Tamhane (1987), Bechofer et. al (1995), Futschick y Pflug (1995, 1997) y Nelson et. al (2001). Kleywegt igualmente cita otra metodología de aproximación a la solución del problema, que consiste en realizar ciertas modificaciones a la meta heurística Temple Simulado, de modo que se use su estructura general, pero reconociendo que la función de optimización tiene términos que pueden cambiar entre iteraciones, debido a su naturaleza probabilística. Como ejemplo de tal metodología de aproximación a la solución del problema se cita a Gelfand y Mitter (1989), Alrefaei y Abdradóttir (1999), Fox y Heine (1995), Gutjahr y Pflug (1996) y Homem de Mello (1999). Por otra parte, Birge y Loveaux (1997) han propuesto un esquema de solución jerárquico para problemas de programación estocásticos de dos etapas. Norkin et. al (1998) sugiere un procedimiento de ramificación y acotamiento para la solución de un modelo de programación matemática con variables

enteras. Por último, en la misma línea que el presente trabajo, esto es, una aproximación a la solución de un problema de programación con variables enteras usando simulación Monte Carlo, se puede citar el trabajo de Morton y Wood (1998).

En cuanto a trabajos de diseño de redes de suministro que han generado la preocupación por un abordaje probabilístico en los parámetros de las redes, puede citarse los trabajos de Escudero et. al (1999, 2002), Vidal y Goetschalckx (2000), Birge (2002), Dupacova (2002), Dormer et. al (2003), Hagle (2003), Mitra et. al (2004), Parija et. al (2004) y Santoso et. al (2005).

### 3. Estrategia algorítmica de solución

Reformulemos el problema que se desea resolver, en un modo más conveniente para propósitos del diseño del algoritmo de solución. Considérese entonces la siguiente formalización:

$$\min_y f(y) = cy + E[Q(y, \xi)] \quad (2)$$

$$\text{Sujeto a: } y \in Y \subseteq \{0,1\}^{|P|} \quad (3)$$

En el problema (2) – (3),  $c$  es un vector de costos asociados a decisiones de primer nivel, como apertura de instalaciones, uso de máquinas, decisiones de producción de producto o similares. A su vez, el vector  $y$  es un vector de variables binarias relacionado con la toma de decisiones en ese primer nivel. Nótese que la  $|P|$  que se incluye en la definición del dominio para el vector  $y$ , está relacionada con el tamaño del problema, en términos del número de decisiones de primer nivel que serían abordadas.  $Q(y, \xi)$  hace referencia a al valor óptimo del siguiente problema de optimización:

$$\min_x q^x \quad (4)$$

$$\text{Sujeto a: } \begin{aligned} g_j(x) &\leq b_j \\ R(x) &\leq M(y) \end{aligned} \quad (5)$$

En el problema (4) – (5)  $x$  es un vector de variables de segundo nivel, como decisiones de cantidad de producto a producir, tamaño de flujos entre nodos, entre otros. El vector  $q$  se compone de los coeficientes de costo asociados a las variables del vector  $x$ . El primer conjunto de restricciones en (5) hace referencia a aspectos que tienen relación con el funcionamiento del sistema una vez que se ha

tomado una cierta decisión en el primer nivel. El segundo conjunto de restricciones sirve de enlace entre las decisiones de primer nivel, representadas por el vector  $y$ , y las de segundo nivel, representadas por el vector  $x$ . Por ejemplo, si una decisión de primer nivel es NO producir un cierto producto en una instalación industrial, una restricción del segundo conjunto impedirá, por su propia estructura matemática, que se asigne una cantidad de producción del mismo en dicha instalación.

En la ecuación (2),  $\xi$  es un vector cuyos elementos son aleatorios, y representan la componente probabilística del problema de optimización en todos los parámetros que así lo requieran. El valor óptimo del problema (4) – (5) es una función de las variables de primer nivel  $y$ , y una realización particular, o escenario, del vector aleatorio  $\xi$ . El valor esperado en la ecuación (2) se toma con respecto a la distribución de probabilidad del vector aleatorio  $\xi$ , lo cual supone que se conoce dicha distribución.

El modelo matemático descrito por las ecuaciones (2) – (5) representa un problema de programación matemática de dos etapas, con elementos aleatorios. La primera etapa representa decisiones de configuración, mientras la segunda representa decisiones de procesamiento y de flujo, que serán tomadas teniendo en cuenta las decisiones de primer nivel, y un escenario particular de los parámetros aleatorios. El objetivo es minimizar los costos presentes de la inversión, representados en las decisiones de primer nivel que de hecho se ejecutarían de inmediato, más los costos futuros esperados, representados por la expresión  $E[Q(y, \xi)]$ .

Se puede citar dos fuentes potenciales de dificultad en la solución del problema (2) – (3): primero, una evaluación de la función objetivo  $f(y)$  para una configuración de primer nivel dada, supone el cálculo del valor esperado de la función lineal  $Q(y, \xi)$ . Sin embargo, recordando la naturaleza aleatoria del vector  $\xi$ , puede suceder que las funciones de probabilidad sean de naturaleza continua, en cuyo caso habría que calcular varias integrales para poder obtener un resultado analítico del valor esperado de la función  $Q$ , proceso de integración que puede resultar prácticamente imposible. Por otra parte, si suponemos que las distribuciones de probabilidad son discretas, el cálculo del valor esperado supondría resolver un número elevado de subproblemas (4) – (5), uno por cada escenario que resulte de la combinación de los valores discretos que pueden tomar los parámetros.

El número de tales subproblemas es usualmente prohibitivo desde una perspectiva computacional; segundo, aún en los casos en que el valor esperado pueda obtenerse de modo exacto, la optimización de la función objetivo dada por la ecuación (2) presenta dificultades. El término  $E[Q(y, \xi)]$  es una función convexa no lineal del vector de variables de primer nivel  $y$ . Tal función no está definida, en todo caso, de un modo analítico directo, pues supone simplemente una suma lineal de distribuciones de probabilidad conocidas, pero arbitrarias en un sentido general. El problema en (2) – (3) supone entonces la minimización de una función apenas implícitamente definida, de tipo no lineal, con respecto a un conjunto de variables binarias. Tal proceso de optimización resulta a menudo altamente complejo.

### 3.1 Sample Average Approximation - SAA

La metodología SAA aborda el problema de cálculo del valor esperado  $E[Q(y, \xi)]$  mediante la generación de muestras del vector aleatorio  $\xi$ , y el posterior proceso de promediar los resultados. Se genera una muestra aleatoria  $\xi^1, \dots, \xi^N$  de  $N$  escenarios del vector  $\xi$ , de modo que el valor esperado  $E[Q(y, \xi)]$  es aproximado por la función promedio de las muestras obtenidas. Así, el problema original dado por las ecuaciones (2) – (3) es aproximado por el problema:

$$\min_{y \in Y} \left\{ \hat{f}_N(y) = cy + \frac{1}{N} \sum_{n=1}^N Q(y, \xi^n) \right\} \quad (6)$$

Sean  $v_N$  y  $\hat{y}_N$  el valor óptimo y el vector de la solución óptima respectivamente, para el problema SAA que representa la ecuación (6). Es de notar que tanto  $v_N$  como  $\hat{y}_N$  son aleatorios, en el sentido que son funciones de la correspondiente muestra aleatoria de tamaño  $N$ . Sin embargo, para una realización particular de la muestra  $\xi^1, \dots, \xi^N$ , el problema de optimización (6) es determinístico y puede ser solucionado por las técnicas adecuadas de optimización clásica.

Kleywegt (2001) ha demostrado que, en tanto el tamaño de la muestra  $N$  se aumenta,  $v_N$  y  $\hat{y}_N$  convergen con probabilidad uno al valor de sus contrapartes en el problema original, y más aún, que  $\hat{y}_N$  converge a la solución óptima del problema original con probabilidad que se aproxima a uno de modo exponencial respecto al tamaño de la muestra. El análisis de convergencia sugiere que una aproximación medianamente buena a la solución













