International Journal of Mathematics and
Computer Science, **19**(2024), no. 1, 157–162

# Using Alternating Direction Method of Multipliers to solve optimization problem in Statistics

**Ameer Dehyauldeen A. Al-Zamili, Ahmed Sabah Ahmed Aljilawi**

Department of Mathematics
Faculty of Education College for Pure Sciences
Babylon University
Babylon, Iraq

email: almosawyameer@gmail.com, aljelawy2000@yahoo.com

(Received July 1, 2023, Accepted August 1, 2023,
Published August 31, 2023)

**Abstract**

Due to its success in handling large optimization problems, the iterative alternating direction method of multipliers (ADMM) has garnered attention. In this article, we discuss the ADMM algorithm's history, theoretical characteristics, and applications. Moreover, we explore new ADMM improvements and their benefits over alternative optimization approaches. Furthermore, we discuss ADMM research gaps and future prospects.

## 1 Introduction

Mathematical models help optimize possible solutions. A model optimizes a constrained objective function for engineering, resource allocation, logistics, scheduling, and portfolio management [5].
Statistics collects, analyzes, interprets, summarizes, and organizes data. It concludes, predicts, and infers. Statistics quantify variability, understands ambiguity, and supports decision-making with limited facts.[5].

**Key words and phrases:** Mathematical model, Convex, ADMM, linear regression, polynomial regression,image processing.
**AMS (MOS) Subject Classifications**: 49M05, 65K10.
**ISSN** 1814-0432, 2024, http://ijmcs.future-in-tech.net

Statistics estimate and assess model parameters and assumptions in optimization. Statistical analysis reveals trends and informs objective functions and limitations. Optimization optimizes statistical estimates and data collecting[7].
Data science combines optimization and statistics. Big data provides useful insights. Machine learning improves model parameters. Regression analysis, hypothesis testing, and Bayesian inference analyze data, verify models, and predicts[5].

**Definition 1.1.** *[1] Convex Set: If the line segment between any two points in C, lies in C. Then the set C is Convex Set*

$$(\theta X_1 + (1-\theta)X_2 \in C), \forall (X_1, X_2) \in C, \forall \theta \in [0,1] \tag{1.1}$$

**Definition 1.2.** *[5] Convex Funtion: A function $f : R^n \to R$ is Convex, if for every $x_1, x_2 \in R_n, 0 \le t \le 1$ the following inequality is true:*

$$f(tx_1 + (1-t)x_2) \le tf(x_1) + (1-t)f(x_2). \tag{1.2}$$

# 2 ADMM Algorithm:

ADMM is an iterative algorithm that minimizes a composite objective function of the form:

$$\begin{array}{ll} \min_{\mathrm{x,z}} & \mathrm{p(x)+q(z)} \\ \mathrm{s.t.} & \mathrm{Mx+Nz=a,} \end{array} \tag{2.3}$$

where $p(x)$ and $q(z)$ are the optimization variables, $p$ and $q$ are convex functions, $M$ and $N$ are the constraint matrices, and $a$ is the constraint vector [4].
The ADMM algorithm operates by introducing a supplementary variable u and transforming the original problem into a problem of equivalent form:

$$\begin{array}{ll} \min_{x,z,u} & f(x)+g(z) \\ \mathrm{s.t} & Ax+Bz=c, x-z=u \end{array} \tag{2.4}$$

The ADMM algorithm then iteratively resolves the subsequent subproblems [2]:

1. Update $x$ by minimizing the augmented Lagrangian function:

$$\mathrm{x^{k+1}} = \arg\min_{\mathrm{x}} \left[\mathrm{f(x)} + \frac{\rho}{2}\left\|\mathrm{x-z^k+u^k}\right\|_2^2\right] \tag{2.5}$$

2. Update $z$ by minimizing the augmented Lagrangian function:

$$\mathrm{z^{k+1}} = \arg\min_{\mathrm{z}} \left[\mathrm{g(z)} + \frac{\rho}{2}\left\|\mathrm{x^{k+1}} - \mathrm{z} + \mathrm{u^k}\right\|_2^2\right] \tag{2.6}$$

3. Update the dual variable $u$ by

$$\mathrm{u^{k+1}} = \mathrm{u^k} + \mathrm{x^{k+1}} - \mathrm{z^{k+1}}, \tag{2.7}$$

where $\rho$ is a positive scalar parameter called the penalty parameter, and $k$ is the iteration number.

## 2.1 A simple example

Let us use the Alternating Direction Method of Multipliers (ADMM) to solve the following problem by hand:

$$\begin{aligned} &Minimize : f(x, y) = x^2 + y^2 \\ &Subjectto : x + y = 1 \end{aligned} \tag{2.8}$$

Step 1: Reformulate the problem using augmented Lagrangian. The augmented Lagrangian for this problem is given by:

$$L_\gamma(x, y, u) = f(x, y) + u(x + y - 1) + \gamma/2(x + y - 1)^2, \tag{2.9}$$

where $x$ and $y$ are the variables, $u$ is the dual variable (Lagrange multiplier), and $\gamma$ is a positive penalty parameter.

Step 2: initialize variables Set an initial guess for $x, y$, and $u$. Let's start with $x = 0, y = 0$, and $u = 0$.

Step 3: Iterate until convergence
Repeat the following steps until convergence is achieved:

a) Update $x$ :

$$x(k+1) = argmin_x x^2 + (\gamma/2) * (x + y(k) - 1 + u(k)/\gamma)^2 \tag{2.10}$$

b) Update $y$ :

$$y(k+1) = argmin_y y^2 + (\gamma/2) * (x(k+1) + y - 1 + u(k)/\gamma)^2 \tag{2.11}$$

c) Update $u$ :

$$u(k+1) = u(k) + \gamma * (x(k+1) + y(k+1) - 1) \tag{2.12}$$

Step 4: Check for convergence:
Repeat steps 3 until the convergence criteria are met.

It will iterate and update $x, y$, and $u$ until convergence is reached. We may observe convergence at iteration No. 56.

# 3 Mathematical Model

In image processing, Gaussian models are often used to represent and analyze the statistical properties of pixel intensities in an image. The Gaussian distribution, also known as the normal distribution, is a continuous probability distribution that is characterized by its mean ($\mu$) and variance ($\sigma^2$) [12].
The Gaussian mathematical paradigm for image processing presupposes that pixel intensities are Gaussian. Natural pictures have consistent intensity changes and noise may be represented as random perturbations around the genuine pixel values. [12].
Mathematically, the Gaussian distribution is defined by the probability density function (PDF):

$$f(x) = \left(1/\sqrt{(2\pi\sigma^2)}\right) \exp\left(-(x-\mu)^2/\left(2\sigma^2\right)\right), \tag{3.13}$$

where $x$ is the pixel intensity value, $\mu$ is the mean of the distribution which represents the average intensity value of the image, and $\sigma^2$ is the variance of the distribution, which measures the spread or variability of the pixel intensities

# 4 Applications

Compressed sensing, image processing, and machine learning use ADMM. ADMM recovers sparse signals from compressed sensing undersampled observations. The ADMM algorithm denoises and deblurs pictures. Support vector machines and logistic regression are optimized using the ADMM technique [1].

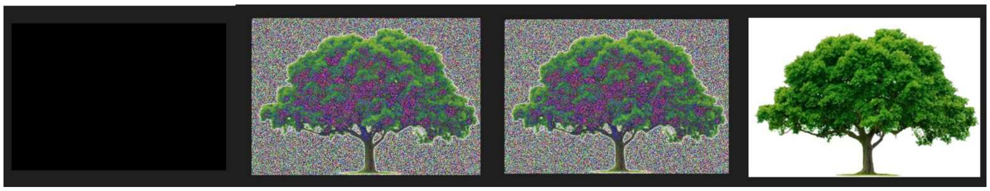

Figure 1: The preceding images illustrate the operation of the ADMM algorithm in image processing.

Here is an example of using the ADMM algorithm in Python for image denoising and deblurring using the **cvxpy** and **numpy** packages.

This example loads a picture, adds Gaussian noise and blur, and creates a noisy-blurred image. We define ADMM-based denoising and deblurring functions. Each function returns the denoised or deblurred picture from the input and noisy-blurred image.
We then denoise and deblur the input picture and show the original, noisy-blurred, denoised, and deblurred images using $P$.

# 5 Conclusions

In this study, we gave an overview of the Alternating Direction Method of Multipliers algorithm, examining its history, theoretical features, practical applications, current advances, and future prospects. ADMM solves complex, large-scale issues through iterative optimization. Scalability, versatility, and systematic issue handling have made it useful in compressed sensing, image processing, and machine learning. Recent breakthroughs address particular difficulties with algorithmic enhancements, theoretical advances, and expansions. The study underlines ADMM's relevance in numerous fields and suggests further research and applications.

# References

[1] Stephen Boyd et al., ”Distributed optimization and statistical learning via the alternating direction method of multipliers,” Foundations and Trends in Machine learning, **3,** no. 1, (2011), 1–122.

[2] Chengbo Li et al., ”An efficient augmented Lagrangian method with applications to total variation minimization,” Computational Optimization and Applications, **56,** (2013), 507–530.

[3] Jianqiang Hu et al., ”Hierarchical interactive demand response power profile tracking optimization and control of multiple EV aggregators,” Electric Power Systems Research, **208,** (2022), 107894.

[4] G. Steidl, T. Teuber, “Removing multiplicative noise by Douglas-Rachford splitting methods,” Journal of Mathematical Imaging and Vision, **36,** no. 2, (2010), 168–184.

[5] Stephen P. Boyd, Lieven Vandenberghe, Convex optimization, Cambridge university press, 2004.

[6] James, Gareth et al., An introduction to statistical learning, **112,** Springer, New York, 2013.

[7] Gerard Cornuejols, Reha Tütüncü, Optimization methods in finance, **5,** Cambridge University Press, 2006.

[8] Jonathan F. Bard, Practical bilevel optimization: algorithms and applications, **30,** Springer Science Business Media, 2013.

[9] Marc Schoenauer, Spyros Xanthakis, Constrained ga optimization. In Proc. of the 5th International Conference on Genetic Algorithms, Morgan Kaufmann, (1993), 573–580.

[10] J. Sun,L. W. Zhang, Y. Wu, ”Properties of the Augmented Lagrangian in Nonlinear Semidefinite Optimization,” Journal of Optimization Theory and Applications, **129,** (2006), 437–456.

[11] Jagdish S. Rustagi, Optimization Techniques in Statistics. Elsevier, 2014.

[12] Tyler Seacrest, ”Mathematical models of image processing,” (2006).