



Theoretical Framework

The Landweber iteration is an iterative regularization method for solving ill-posed or inconsistent linear systems of the form: $X\beta = y$, where X is a $n \times m$ design matrix, y is an $n \times 1$ response vector and β is an $m \times 1$ parameter vector we want to estimate [1], [2]. The method iteratively updates β to reduce the error. It leverages the properties of Banach spaces to handle non-smooth solutions and promote sparsity, which is crucial for variable selection in high-dimensional statistical models [3], [4]. By using duality mappings, the algorithm effectively generalizes the Landweber method from Hilbert to Banach spaces, allowing for regularization in spaces where traditional least squares methods fail.

Algorithm

Algorithm 1 $J_duality_map(\beta, p, s)$

```
1: if  $\|\beta\| \neq 0$  then
2:    $R \leftarrow (\sum_{i=1}^n |\beta_i|^p)^{s/p-1} \cdot (|\beta|^{p-1} \times \text{sign}(\beta))$ 
3: else
4:    $R \leftarrow \beta$ 
5: end if
6: return  $R$ 
```

Algorithm 2 $\text{Landweber}(X, y, MAXIT, p, tol)$

```
1:  $n, m, q, s, r, \mu \leftarrow$  rows of  $X$ , columns of  $X$ ,  $\frac{p-1}{p}, 2, \frac{s}{s-1}, \frac{1}{\|A\|^2}$ 
2:  $\beta_0, \beta_{\text{new}} \leftarrow 0_{n \times 1}, 0_{n \times 1}$ 
3: while  $i \leq MAXIT$  and  $|\beta_{\text{new}} - \beta_{\text{old}}| \leq tol$  do
4:    $\beta_{\text{old}} \leftarrow \beta_{\text{new}}$ 
5:    $\text{res} \leftarrow X \cdot \beta_{\text{old}} - y$ 
6:    $R \leftarrow J\_duality\_map(\text{res}, p, s)$ 
7:    $\beta_{\text{new}} \leftarrow J\_duality\_map(J\_duality\_map(\beta_{\text{old}}, p, s) - \mu \cdot X^T \cdot R, q, r)$ 
8: end while
9: return  $\beta_{\text{new}}$ 
```

Stop Criteria

The algorithm stops when a predefined maximum number of iterations is reached or when the relative change in the coefficient vector β falls below a certain threshold, $\frac{\|\beta_{k+1} - \beta_k\|}{\|\beta_k\|} < tol$, where tol is a small tolerance, selected based on the application.

Algorithm's complexity

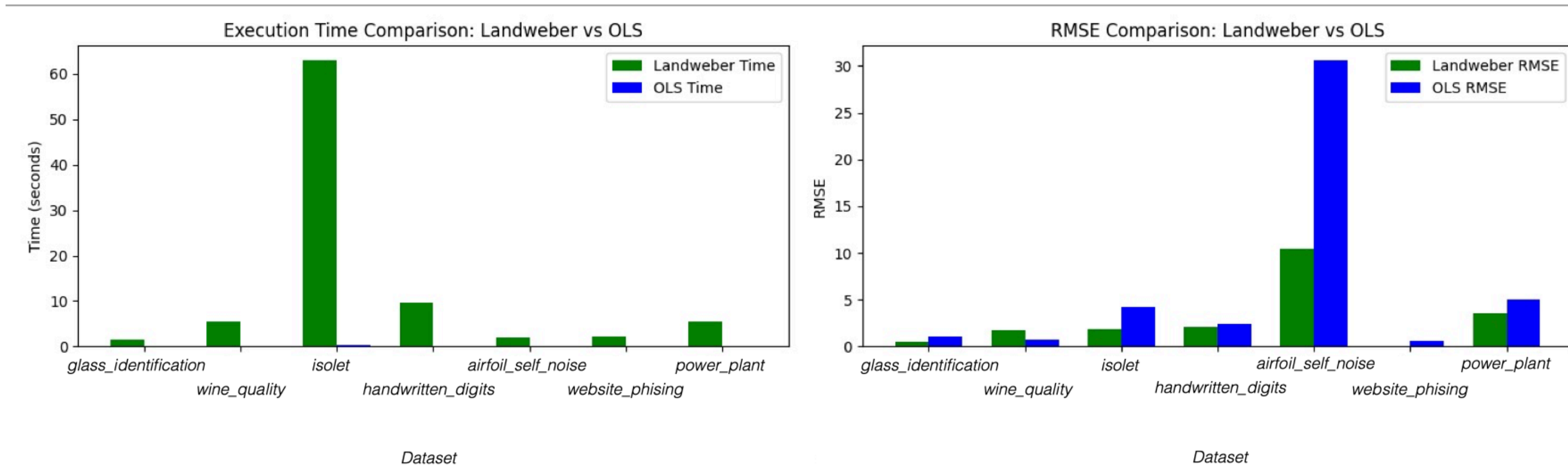
Let us that the design matrix $X \in \mathbb{R}^{n \times m}$. Then $\beta \in \mathbb{R}^{m \times 1}$ and $y \in \mathbb{R}^{n \times 1}$. Firstly, the computational complexity of the duality map is linear with respect to its input, therefore $O(m)$. The computational complexity per Landweber iteration is primarily determined by matrix-vector multiplications and duality mappings. The computation of the residual, $X\beta - y$ is $O(mn)$, the duality mapping applied to it is of linear complexity with respect to its input, therefore line 6 has complexity $O(mn)$. The same holds true for line 7. The first duality map is applied with input of dimensions $m \times 1$ and is of complexity $O(m)$, the $\mu \cdot X^T \cdot R$ is of complexity $O(mn)$, therefore $J_duality_map(J_duality_map(\beta_{\text{old}}, p, s) - \mu \cdot X^T \cdot R, q, r)$ is of complexity $O(mn)$.

For a maximum of MAXIT iterations, the worst-case complexity of the algorithm is $O(MAXIT \cdot n \cdot m)$. Early stopping based on the convergence of the root mean square error (RMSE) can reduce the effective number of iterations k , leading to an improved practical complexity of $O(k \cdot n \cdot m)$.

Real World Datasets

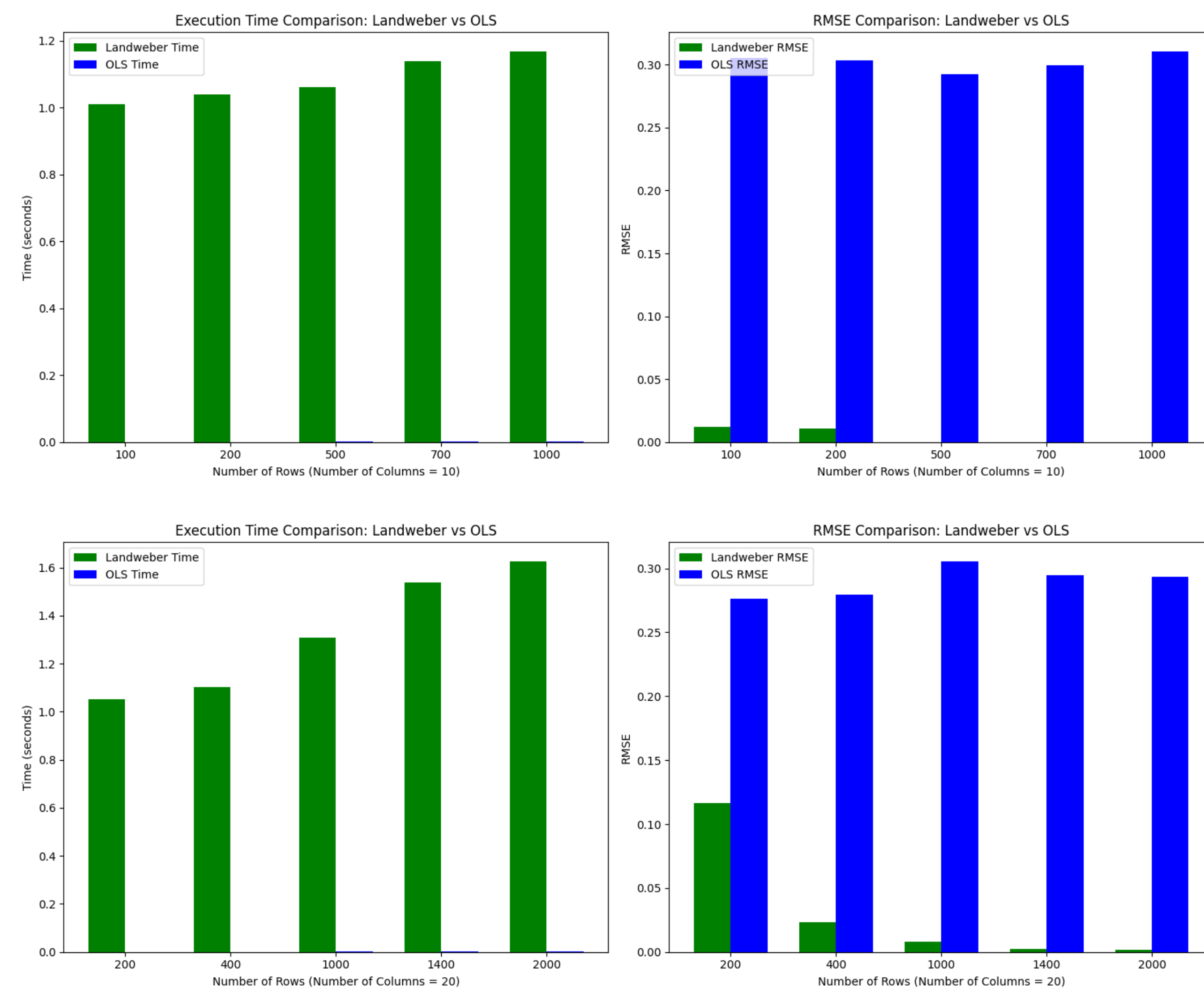
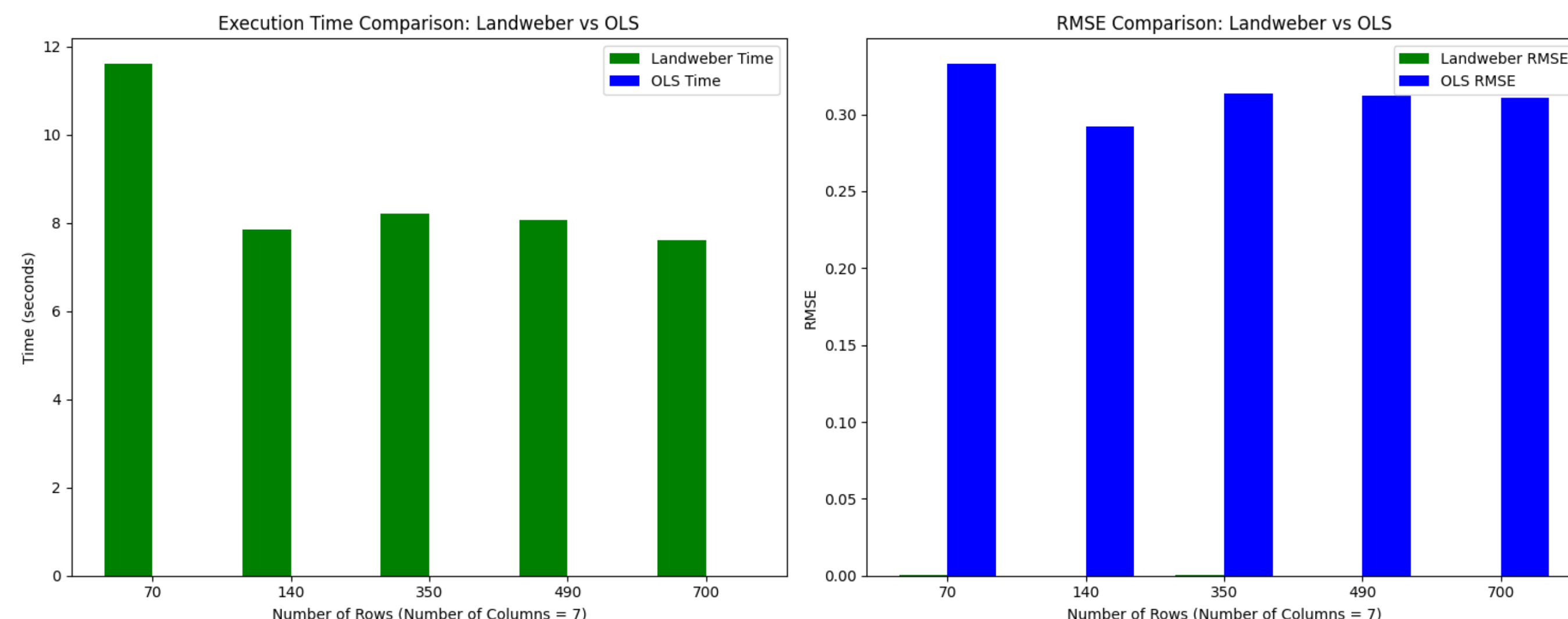
The algorithm was tested on several real-world datasets. The Landweber iterative method was applied to solve the linear system $y = X\beta + \varepsilon$. All datasets exhibited highly overdetermined design matrices, reflecting real-world scenarios where observations vastly outnumber predictors. Uniform noise ($\sigma = 0.1$) was introduced to the target variable y to simulate real-world imperfections. The method ran for a maximum of 5000 iterations, ensuring convergence within a tolerance of 10^{-6} .

- Glass Identification Dataset
- Wine Quality Dataset
- Isolet Dataset
- Pen-Based Recognition of Handwritten Digits
- Airfoil Self Noise
- Website Phishing
- Combined Cycle Power Plant



Simulated Data Benchmarks

For further evaluation of the Landweber method, simulated data was generated with varying matrix dimensions, testing the performance across different configurations. The design matrices X and response vectors y were constructed using a uniform distribution, with column sizes fixed at 5, 7, 10, and 20, and row sizes scaled as multiples (10, 20, 50, 70, 100) of the columns, resulting in overdetermined systems (e.g., 50rows \times 5 columns, up to 2000 \times 20 columns). Uniform noise ($\sigma = 0.1$) was added to the response vector. The Landweber method was run for a maximum of 500 iterations, optimizing the parameter p in the range [1.1, 2.1] to minimize RMSE, while OLS provided a baseline for comparison.



Conclusions

The results reveal that Landweber consistently outperforms OLS in terms of RMSE across all simulated configurations in terms of RMSE. However, Landweber's execution time is significantly higher, often exceeding 10 seconds, compared to OLS, which completes in under 0.01 seconds. This trend holds across both real and simulated data, highlighting Landweber's key advantage: its ability to achieve superior accuracy, making it ideal for applications where precision is critical, such as predictive maintenance in power plants or phishing detection. However, the method's computational cost poses a drawback, particularly for large-scale or time-sensitive applications. Overall, the Landweber method offers a compelling trade-off for scenarios prioritizing accuracy over speed, but its scalability remains a challenge for real-time systems.

References

- [1] Estatico, C., Koukouvinos, C., Mitrouli, M., & Tsagkaropoulos, I. (2024, October). Landweber Banach regression in statistical learning [Poster presentation]. The 14th International Conference on Dependable Systems, Services and Technologies (DESSERT2024), Athens, Greece.
- [2] Brianzi, P., Di Benedetto, F., & Estatico, C. (2013). Preconditioned iterative regularization in Banach spaces. Computational Optimization and Applications, 54(2), 263–282. <https://doi.org/10.1007/s10589-012-9527-2>
- [3] Georgiou, S.D. Modelling by supersaturated designs. Comput. Stat. Data Anal. 2008, 53, 428–435.
- [4] Koukoudakis, N., Koukouvinos, C., Lappa, A., Mitrouli, M., Psitou, A. Numerical methods in modeling with supersaturated designs, Appl. Numer. Math, (to appear) <https://doi.org/10.1016/j.apnum.2024.02.003>.