

Dear readers,

Eq.(10) can have a different version of the Kronecker product.

According to Theorem 2.2 in [1], the following should hold:

$$\frac{\partial \text{vec}(aA(x))}{\partial x} = (I \otimes a) \frac{\partial \text{vec}(A(x))}{\partial x} \quad (1)$$

where a is a vector, A is a matrix depending on a vector x . In this case, the order of terms in the Kronecker product of Eq.(10) should be reversed.

However, my simulation results show that strictly following this definition can cause numerical issues, like complex eigenvalues. Therefore, we opted to use the definition in Eq.(4) of [2] instead. Our algorithm can still work using definition (1) by taking only the real parts of eigenvalues, although this requires fine tuning of the hyperparameters of the trust-region method.

This is an interesting phenomenon worth further investigation.

[1]. Magnus, Jan R., and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons, 2019.

[2]. Dyro, Robert, Edward Schmerling, Nikos Arechiga, and Marco Pavone. "Second-Order Sensitivity Analysis for Bilevel Optimization." In *International Conference on Artificial Intelligence and Statistics*, pp. 9166-9181. PMLR, 2022.

王秉亨

WANG Bingheng

31-Oct-2024