Let's Make Block Coordinate Descent Converge Faster
Julie Nutini, Issam Laradji, Mark Schmidt

8 Ways to Make BCD Converge Faster*

1. Use fixed blocks with random rules.
2. Use greedy rules to pick the block instead of cyclic/random.
3. Use variable blocks with greedy rules.
4. Update using Newton’s method and a line search.
5. Use Lipschitz information with greedy rules.
6. Increase the block size.
7. Use tree-structured blocks if you have sparse dependencies.
8. Use two-metric projected Newton for non-smooth problems.

*If it does not significantly increase the iteration cost.

Ask us about superlinear convergence under the right conditions!

Experimental Results

- Random faster with fixed blocks.
- Greedy faster than random.
- Greedy faster with variable blocks.
- Newton updates faster than matrix.
- Greedy-Lipschitz faster than greedy.
- Bigger blocks converge faster.
- Colouring faster than using small blocks.
- Trees converge faster than colouring.
- Projected-Newton faster than gradient.
- TMP similar to Newton with low cost.
- Identifies active set in finite time.

Scan for our codebase