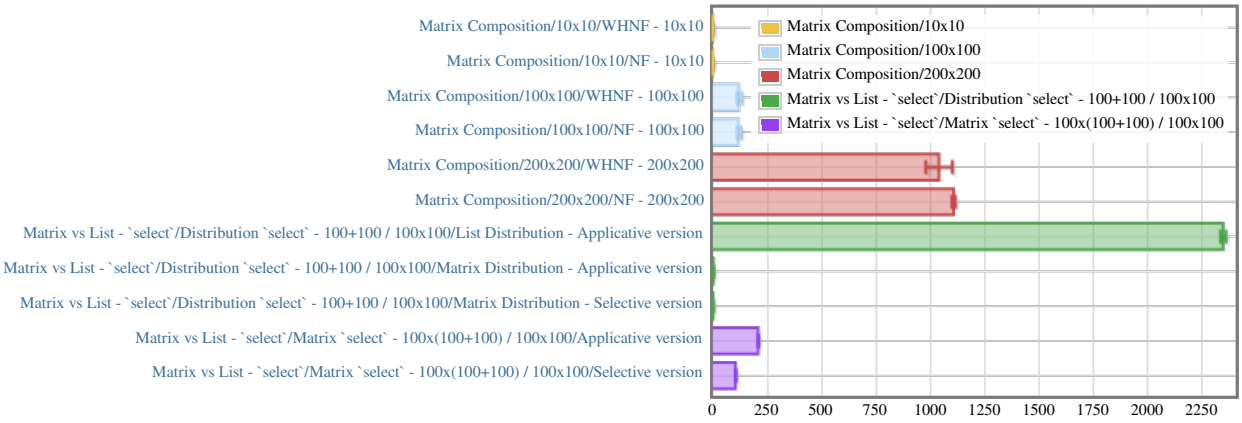
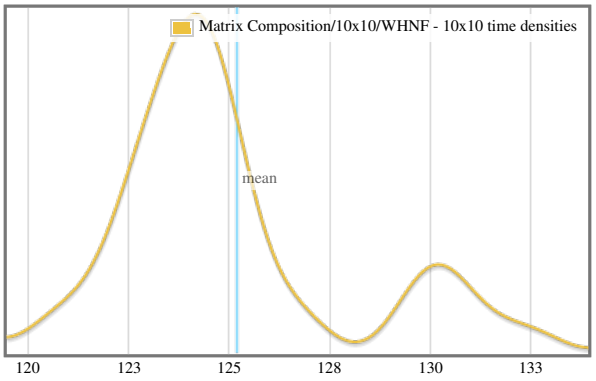
criterion performance measurements

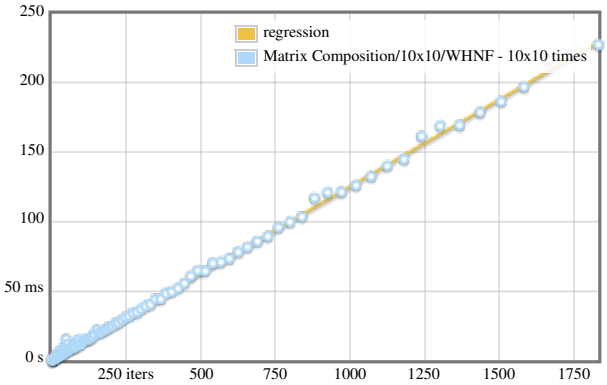
overview

want to understand this report?



Matrix Composition/10x10/WHNF - 10x10

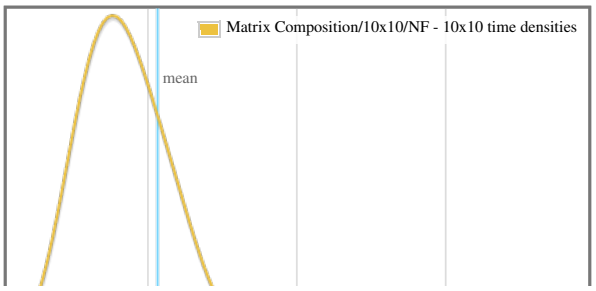


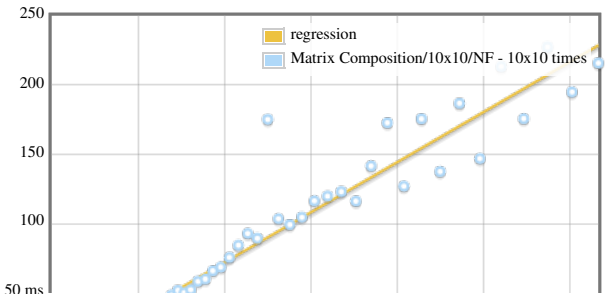


	lower bound	estimate	upper bound
OLS regression	124 µs	125 µs	127 µs
R² goodness-of-fit	0.998	0.999	0.999
Mean execution time	124 µs	125 µs	126 µs
Standard deviation	2.38 µs	3.00 µs	3.87 µs

Outlying measurements have moderate (19.5%) effect on estimated standard deviation.

Matrix Composition/10x10/NF - 10x10

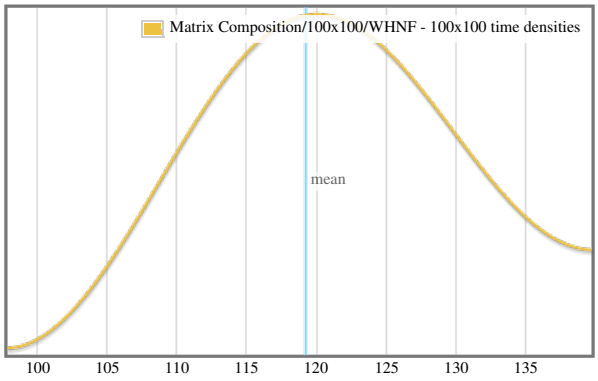


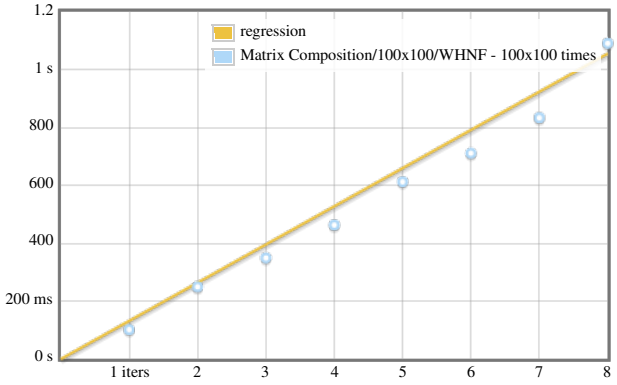


	lower bound	estimate	upper bound
OLS regression	135 µs	144 µs	156 µs
R² goodness-of-fit	0.919	0.960	0.985
Mean execution time	145 µs	153 µs	167 µs
Standard deviation	23.7 µs	36.5 µs	53.6 µs

Outlying measurements have severe (96.1%) effect on estimated standard deviation.

Matrix Composition/100x100/WHNF - 100x100

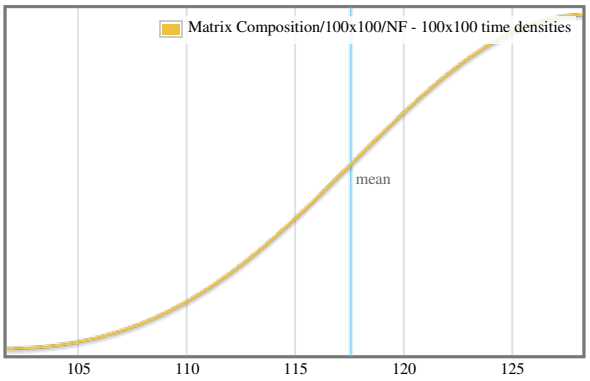


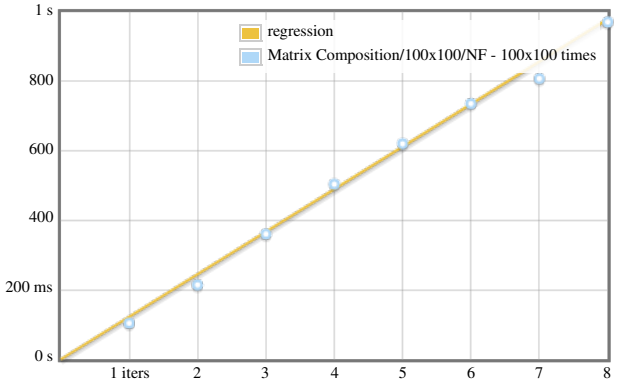


	lower bound	estimate	upper bound
OLS regression	117 ms	132 ms	149 ms
R² goodness-of-fit	0.970	0.985	0.999
Mean execution time	113 ms	119 ms	126 ms
Standard deviation	6.39 ms	9.78 ms	16.2 ms

Outlying measurements have moderate (23.2%) effect on estimated standard deviation.

Matrix Composition/100x100/NF - 100x100

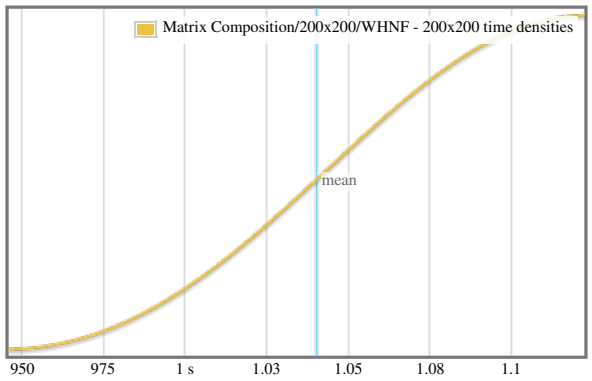


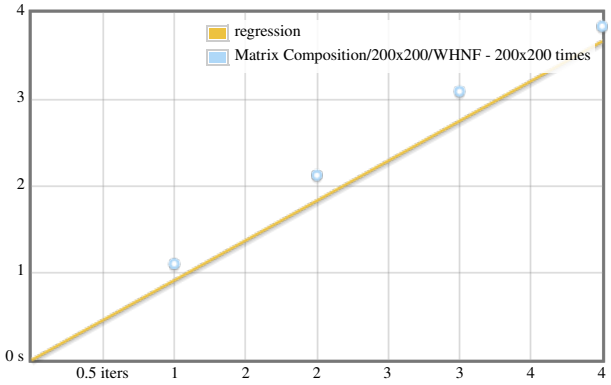


	lower bound	estimate	upper bound
OLS regression	111 ms	122 ms	131 ms
R² goodness-of-fit	0.987	0.995	0.999
Mean execution time	111 ms	118 ms	122 ms
Standard deviation	5.37 ms	8.15 ms	11.1 ms

Outlying measurements have moderate (22.5%) effect on estimated standard deviation.

Matrix Composition/200x200/WHNF - 200x200

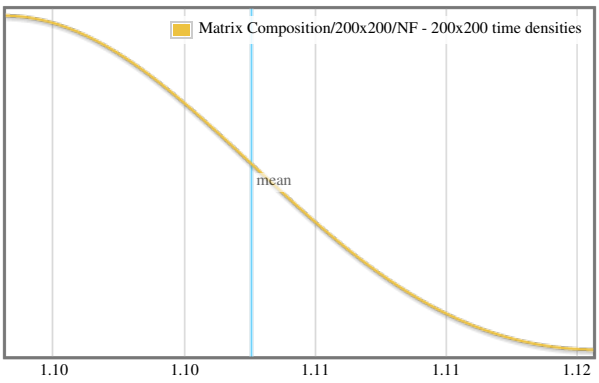


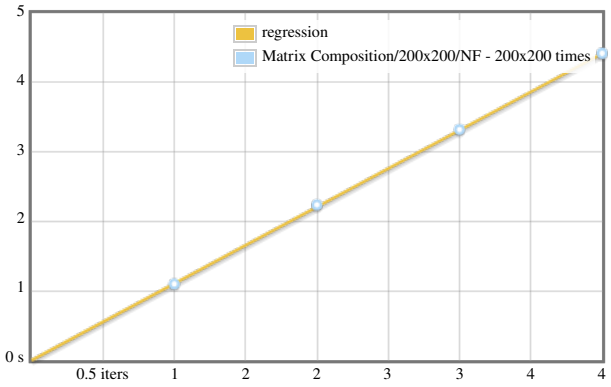


	lower bound	estimate	upper bound
OLS regression	752 ms	917 ms	1.02 s
R² goodness-of-fit	0.992	0.995	1.000
Mean execution time	978 ms	1.04 s	1.09 s
Standard deviation	22.4 ms	61.8 ms	84.8 ms

Outlying measurements have moderate (18.7%) effect on estimated standard deviation.

Matrix Composition/200x200/NF - 200x200

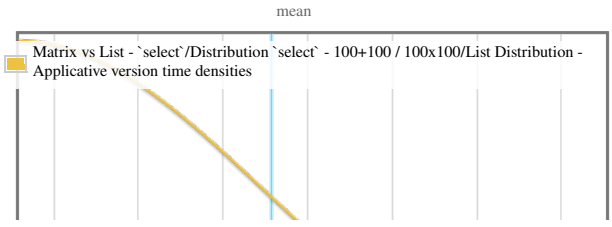


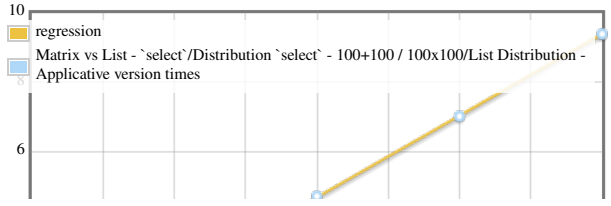


	lower bound	estimate	upper bound
OLS regression	1.08 s	1.10 s	1.14 s
R² goodness-of-fit	1.000	1.000	1.000
Mean execution time	1.10 s	1.11 s	1.12 s
Standard deviation	2.19 ms	7.98 ms	10.8 ms

Outlying measurements have moderate (18.8%) effect on estimated standard deviation.

Matrix vs List - `select`/Distribution `select` - 100+100 / 100x100/List Distribution - Applicative version

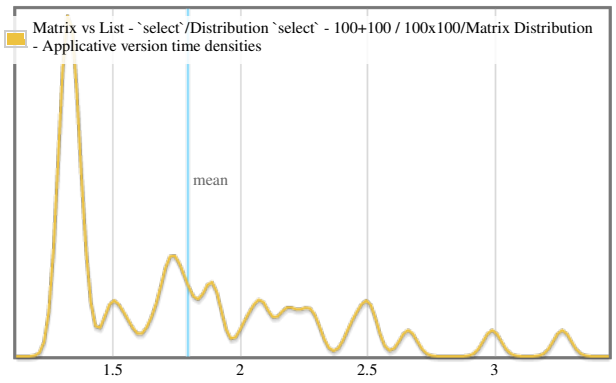


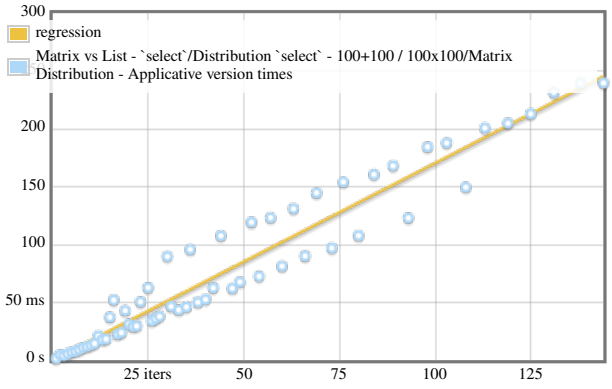


	lower bound	estimate	upper bound
OLS regression	2.30 s	2.35 s	2.37 s
R² goodness-of-fit	1.000	1.000	1.000
Mean execution time	2.34 s	2.35 s	2.36 s
Standard deviation	3.62 ms	12.4 ms	16.7 ms

Outlying measurements have moderate (18.7%) effect on estimated standard deviation.

Matrix vs List - `select`/Distribution `select` - 100+100 / 100x100/Matrix Distribution - Applicative version

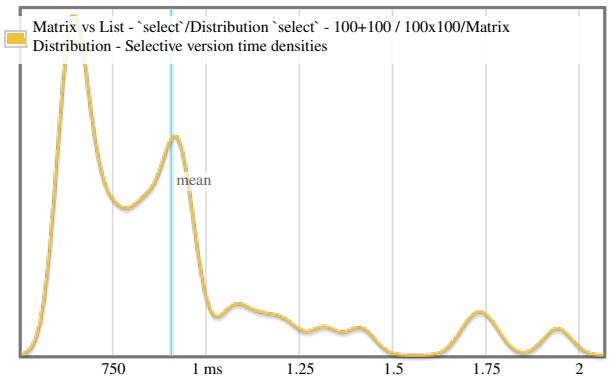


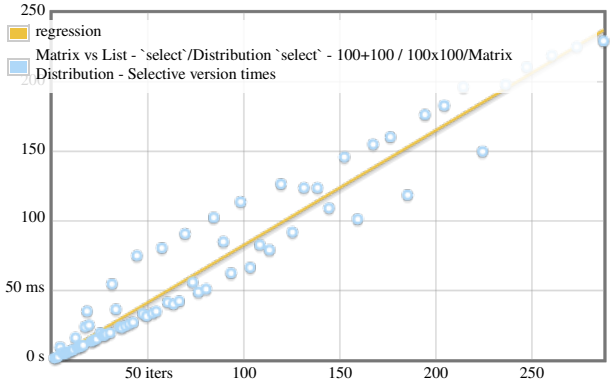


	lower bound	estimate	upper bound
OLS regression	1.61 ms	1.70 ms	1.79 ms
R² goodness-of-fit	0.899	0.939	0.965
Mean execution time	1.64 ms	1.79 ms	1.97 ms
Standard deviation	394 µs	502 µs	659 µs

Outlying measurements have severe (94.7%) effect on estimated standard deviation.

Matrix vs List - `select`/Distribution `select` - 100+100 / 100x100/Matrix Distribution - Selective version



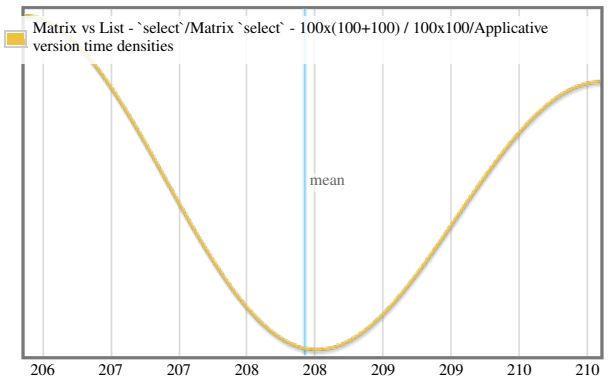


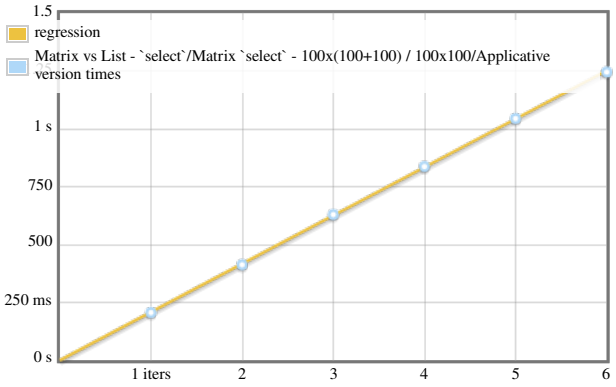
	lower bound	estimate	upper bound
OLS regression	781 µs	828 µs	872 µs

	lower bound	estimate	upper bound
R² goodness-of-fit	0.911	0.947	0.968
Mean execution time	825 µs	904 µs	1.01 ms
Standard deviation	234 µs	318 µs	427 µs

Outlying measurements have severe (98.5%) effect on estimated standard deviation.

Matrix vs List - `select`/Matrix `select` - 100x(100+100) / 100x100/Applicative version

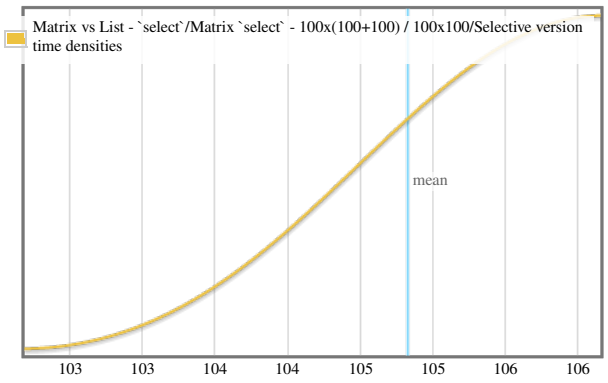


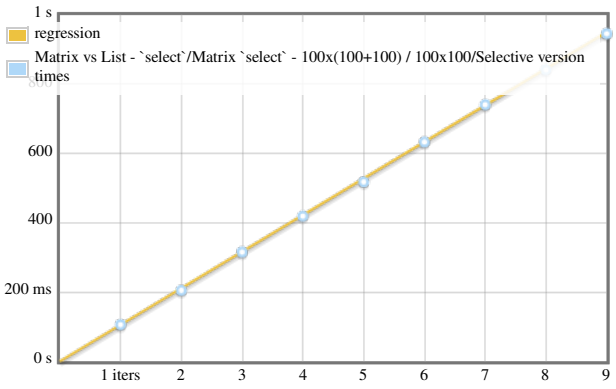


	lower bound	estimate	upper bound
OLS regression	205 ms	208 ms	212 ms
R² goodness-of-fit	1.000	1.000	1.000
Mean execution time	207 ms	208 ms	209 ms
Standard deviation	1.12 ms	1.51 ms	1.83 ms

Outlying measurements have moderate (13.9%) effect on estimated standard deviation.

Matrix vs List - `select`/Matrix `select` - 100x(100+100) / 100x100/Selective version





	lower bound	estimate	upper bound
OLS regression	105 ms	106 ms	107 ms
R² goodness-of-fit	1.000	1.000	1.000
Mean execution time	104 ms	105 ms	105 ms
Standard deviation	586 µs	1.08 ms	1.56 ms

Outlying measurements have slight (9.9%) effect on estimated standard deviation.

understanding this report

In this report, each function benchmarked by criterion is assigned a section of its own. The charts in each section are active; if you hover your mouse over data points and annotations, you will see more details.

- The chart on the left is a kernel density estimate (also known as a KDE) of time measurements. This graphs the probability of any given time measurement occurring. A spike indicates that a measurement of a particular time occurred; its height indicates how often that measurement was repeated.
- The chart on the right is the raw data from which the kernel density estimate is built. The *x* axis indicates the number of loop iterations, while the *y* axis shows measured execution time for the given number of loop iterations. The line behind the values is the linear regression prediction of execution time for a given number of iterations. Ideally, all measurements will be on (or very near) this line.

Under the charts is a small table. The first two rows are the results of a linear regression run on the measurements displayed in the right-hand chart.

- *OLS regression* indicates the time estimated for a single loop iteration using an ordinary least-squares regression model. This number is more accurate than the *mean* estimate below it, as it more effectively eliminates measurement overhead and other constant factors.
- R^2 *goodness-of-fit* is a measure of how accurately the linear regression model fits the observed measurements. If the measurements are not too noisy, R^2 should lie between 0.99 and 1, indicating an excellent fit. If the number is below 0.99, something is confounding the accuracy of the linear model.
- *Mean execution time* and *standard deviation* are statistics calculated from execution time divided by number of iterations.

We use a statistical technique called the bootstrap to provide confidence intervals on our estimates. The bootstrap-derived upper and lower bounds on estimates let you see how accurate we believe those estimates to be. (Hover the mouse over the table headers to see the confidence levels.)

A noisy benchmarking environment can cause some or many measurements to fall far from the mean. These outlying measurements can have a significant inflationary effect on the estimate of the standard deviation. We calculate and display an estimate of the extent to which the standard deviation has been inflated by outliers.

colophon

This report was created using the criterion benchmark execution and performance analysis tool.

Criterion is developed and maintained by Bryan O'Sullivan.