

Principal Component Analysis

Experiment details:

In this section we compare performance of our **PCA** class + **Transform** operation to Matlab's $[W, Y] = \text{princomp}(X)$ function with default settings.

We will use input data in raw format, i.e. data matrix where columns corresponds to objects, rows to their characteristics. Computation time will include metric calculation. So, we have three main parameters for this method:

- Number of objects or number of series (N).
- Number of factors in object characteristic or number of series observations (M).

Performance metrics:

In this test we use the following performance metrics:

1. Total working time – time which was needed to complete specified numbers of computations.
2. FLOPS – approximated number of floating point operations number per second. Typically we use mega flops (MFLOPS which) is $FLOPS \times 10^6$. Denote repeats count as K . So we can write experiment complexity in asymptotic form as $O((N^2M + N^3)K)$. We will use this expression as approximation of floating point calculations count (FLOPS). So for example if we have 250 objects, 10 factors and we make 100 runs in 15 seconds we will say that method performance is:

$$\frac{N^2MK}{\text{computation time}} = \frac{(250^2 * 10 + 250^3) * 100}{15} = 10833333333 \approx 108 \text{ MFLOPS}$$

3. Time of single computation – time which was needed to solve one problem with selected parameters.

Testing system details:

All tests were performed on typical medium class desktop machine.

CPU: Core i7 870 (2.93GHz)
Memory: 8GB RAM
OS: Windows 7 x64

All source code of this test are available, one can get performance results on his own system. Changing test parameters is also possible.

Principal Component Analysis								
Objects	Factors	Repeats	FinMath			Matlab		
			Total Time	MFLOPS	Single Time	Total Time	MFLOPS	Single Time
100	100	50	0.215	466	0.0043	0.257	390	0.005
100	200	50	0.225	667	0.0045	0.402	374	0.008
100	300	50	0.233	859	0.00466	0.474	422	0.009
100	400	50	0.277	903	0.00554	0.558	448	0.011
100	500	50	0.268	1120	0.00536	0.621	483	0.012
500	500	10	1.76	1421	0.176	1.435	1742	0.144
500	1000	10	1.885	1990	0.1885	2.403	1561	0.24
500	1500	10	2.198	2275	0.2198	2.678	1867	0.268
500	2000	10	2.384	2622	0.2384	3.286	1903	0.329
500	2500	10	2.825	2655	0.2825	3.956	1897	0.396
1000	1000	1	1.213	1649	1.213	0.8	2499	0.8
1000	2000	1	1.418	2116	1.418	1.385	2166	1.385
1000	3000	1	1.56	2565	1.56	1.704	2347	1.704
1000	4000	1	1.736	2881	1.736	1.919	2606	1.919
1000	5000	1	1.812	3312	1.812	2.337	2568	2.337

