Notes on GEMV (general matrix vector product)

We only consider the special case where:
- Matrix A is stored row major
- alpha and beta equal one, i.e.

\[ y += Ax \]

We further simplify the cache:
- only 4 cache lines fit in cache
- each cache line can hold 4 elements
- eviction of last used cache lines

Goal: Count operations per fetched cache line
Access pattern: gemv with dot
Cache lines fetched: 3
Mult. operations: 4
Add. operations: 4
Cache lines fetched: 5
Mult. operations: 8
Add. operations: 8
Cache lines fetched: 7
Mult. operations: 12
Add. operations: 12
Access pattern: gemv with fused dot (fuse factor 2)
Cache lines fetched: 4
Mult. operations: 8
Add. operations: 8