WHAT IS CLAIMED IS:

1. A computer-implemented method, comprising:

receiving, with at least one processor, at least one embedding set, each embedding set of the at least one embedding set comprising a set of embedding vectors;

applying, with at least one processor, mean centering to the at least one embedding set;

applying, with at least one processor, spectral normalization to the at least one embedding set, wherein applying spectral normalization to the at least one embedding set comprises:

decomposing, with at least one processor, the at least one embedding set to provide a left singular vector, a right singular vector, and a diagonal matrix;

determining, with at least one processor, an average singular value of the at least one embedding set;

for each respective singular value of the diagonal matrix, determining, with at least one processor, whether the respective singular value is greater than a configurable multiple of the average singular value;

for each respective singular value of the diagonal matrix, if the respective singular value is greater than the configurable multiple of the average singular value, determining, with at least one processor, a respective substitute singular value based on a quotient of the respective singular value divided by the configurable multiple of the average singular value or, if the respective singular value is not greater than the configurable multiple of the average singular value, determining, with at least one processor, the respective substitute singular value to be 1, wherein a substitute diagonal matrix comprises the respective substitute singular value for each respective singular value of the diagonal matrix; and

replacing, with at least one processor, the at least one embedding set with a product of the at least one embedding set, the right singular vector, and an inverse of the substitute diagonal matrix; and

applying, with at least one processor, length normalization to the at least one embedding set, wherein the at least one embedding set comprises a first language embedding set and a second language embedding set, the first language embedding set comprising a first set of word embedding vectors for a first language, the second language embedding set comprising a second set of word embedding vectors for a second language.

2. The method of claim 1, wherein applying mean centering comprises:

determining, with at least one processor, a mean based on all embedding vectors of the set of embedding vectors; and

subtracting, with at least one processor, the mean from each embedding vector of the set of embedding vectors.

3. The method of claim 1, wherein decomposing the at least one embedding set comprises performing singular value decomposition on the at least one embedding set.

4. The method of claim 1, wherein determining the average singular value comprises determining a square root of an average squared singular value.

5. The method of claim 1, wherein applying length normalization comprises:

adjusting, with at least one processor, each embedding vector of the set of embedding vectors to have a 2-norm of 1.

6. The method of claim 1, further comprising:

iteratively repeating, with at least one processor, applying mean centering, applying spectral normalization, and applying length normalization to the at least one embedding set for a configurable number of iterations.

7. The method of claim 1, wherein the at least one embedding set comprises a first embedding set and a second embedding set, the method further comprising:

aligning, with at least one processor, the first embedding set with the second embedding set.

8. A system, comprising:

at least one processor; and

at least one non-transitory computer-readable medium including one or more instructions that, when executed by the at least one processor, direct the at least one processor to:

receive at least one embedding set, each embedding set of the at least one embedding set comprising a set of embedding vectors;

apply mean centering to the at least one embedding set;

apply spectral normalization to the at least one embedding set, wherein applying spectral normalization to the at least one embedding set comprises:

decomposing the at least one embedding set to provide a left singular vector, a right singular vector, and a diagonal matrix;

determining an average singular value of the at least one embedding set;

for each respective singular value of the diagonal matrix, determining whether the respective singular value is greater than a configurable multiple of the average singular value;

for each respective singular value of the diagonal matrix, if the respective singular value is greater than the configurable multiple of the average singular value, determining a respective substitute singular value based on a quotient of the respective singular value divided by the configurable multiple of the average singular value or, if the respective singular value is not greater than the configurable multiple of the average singular value, determining the respective substitute singular value to be 1, wherein a substitute diagonal matrix comprises the respective substitute singular value for each respective singular value of the diagonal matrix; and

replacing the at least one embedding set with a product of the at least one embedding set, the right singular vector, and an inverse of the substitute diagonal matrix; and

apply length normalization to the at least one embedding set, wherein the at least one embedding set comprises a first language embedding set and a second language embedding set, the first language embedding set comprising a first set of word embedding vectors for a first language, the second language embedding set comprising a second set of word embedding vectors for a second language.

9. The system of claim 8, wherein applying mean centering comprises:

determining a mean based on all embedding vectors of the set of embedding vectors; and

subtracting the mean from each embedding vector of the set of embedding vectors.

10. The system of claim 8, wherein decomposing the at least one embedding set comprises performing singular value decomposition on the at least one embedding set.

11. The system of claim 8, wherein determining the average singular value comprises determining a square root of an average squared singular value.

12. The system of claim 8, wherein applying length normalization comprises:

adjusting each embedding vector of the set of embedding vectors to have a 2-norm of 1.

13. The system of claim 8, wherein the instructions, when executed by the at least one processor, further direct the at least one processor to:

iteratively repeat applying mean centering, applying spectral normalization, and applying length normalization to the at least one embedding set for a configurable number of iterations.

14. The system of claim 8, wherein the at least one embedding set comprises a first embedding set and a second embedding set, wherein the instructions, when executed by the at least one processor, further direct the at least one processor to:

align the first embedding set with the second embedding set.

15. The system of claim 8, wherein the at least one embedding set comprises a first language embedding set and a second language embedding set, the first language embedding set comprising a first set of word embedding vectors for a first language, the second language embedding set comprising a second set of word embedding vectors for a second language.

16. A computer program product comprising at least one non-transitory computer-readable medium including one or more instructions that, when executed by at least one processor, cause the at least one processor to:

receive at least one embedding set, each embedding set of the at least one embedding set comprising a set of embedding vectors;

apply mean centering to the at least one embedding set;

apply spectral normalization to the at least one embedding set, wherein applying spectral normalization to the at least one embedding set comprises:

decomposing the at least one embedding set to provide a left singular vector, a right singular vector, and a diagonal matrix;

determining an average singular value of the at least one embedding set;

for each respective singular value of the diagonal matrix, determining whether the respective singular value is greater than a configurable multiple of the average singular value;

for each respective singular value of the diagonal matrix, if the respective singular value is greater than the configurable multiple of the average singular

value, determining a respective substitute singular value based on a quotient of the respective singular value divided by the configurable multiple of the average singular value or, if the respective singular value is not greater than the configurable multiple of the average singular value, determining the respective substitute singular value to be 1, wherein a substitute diagonal matrix comprises the respective substitute singular value for each respective singular value of the diagonal matrix; and

replacing the at least one embedding set with a product of the at least one embedding set, the right singular vector, and an inverse of the substitute diagonal matrix; and

apply length normalization to the at least one embedding set, wherein the at least one embedding set comprises a first language embedding set and a second language embedding set, the first language embedding set comprising a first set of word embedding vectors for a first language, the second language embedding set comprising a second set of word embedding vectors for a second language.