HyperLearn Documentation

Release 1

Daniel Han-Chen

Jun 19, 2020

Contents

HyperLearn aims to make Machine Learning algorithms run in at least 50% of their original time. Algorithms from Linear Regression to Principal Component Analysis are optimized by using LAPACK, BLAS, and parallelized through Numba.

Some key current achievements of HyperLearn:

1. 70% less time to fit Least Squares / Linear Regression than sklearn + 50% less memory usage
2. 50% less time to fit Non Negative Matrix Factorization than sklearn due to new parallelized algo
3. 40% faster full Euclidean / Cosine distance algorithms
4. 50% less time LSMR iterative least squares
5. New Reconstruction SVD - use SVD to impute missing data! Has .fit AND .transform. Approx 30% better than mean imputation
6. 50% faster Sparse Matrix operations - parallelized
7. RandomizedSVD is now 20 - 30% faster

CHAPTER 1

Example code

Singular Value Decomposition

```
import hyperlearn as hl
U, S, VT = hl.linalg.svd(X)
```

Pseudoinverse of a matrix using Cholesky Decomposition

```
from hyperlearn.linalg import pinvc
inv = pinvc(X)

# check if pinv(X) * X = identity
check = inv.dot(X)
```

QR Decomposition wanting only Q matrix

```
from hyperlearn import linalg
Q = linalg.qr(X, Q_only = True)

# check if Q == Q from full QR
q, r = linalg.qr(X)
```

1.1 hyperlearn

1.1.1 hyperlearn package

1.1.1.1 Submodules

1.1.1.2 hyperlearn.base module

hyperlearn.base.**Numpy**(**args*)

hyperlearn.base.**T**(*X*)

hyperlearn.base.**Tensor**(*X*)

hyperlearn.base.**Tensors**(**args*)

hyperlearn.base.**USE_NUMBA = True**
Type Checks Updated 27/8/2018 ————————————————————————

hyperlearn.base.**array**(*X*)

hyperlearn.base.**cast**(*X*, *dtype*)

hyperlearn.base.**check**(*f*)

hyperlearn.base.**constant**(*X*)

hyperlearn.base.**diag**(*input*, *diagonal=0*, *out=None*) → Tensor

- If `input` is a vector (1-D tensor), then returns a 2-D square tensor with the elements of `input` as the diagonal.
- If `input` is a matrix (2-D tensor), then returns a 1-D tensor with the diagonal elements of `input`.

The argument `diagonal` controls which diagonal to consider:

- If `diagonal` = 0, it is the main diagonal.
- If `diagonal` > 0, it is above the main diagonal.
- If `diagonal` < 0, it is below the main diagonal.

Args: input (Tensor): the input tensor. diagonal (int, optional): the diagonal to consider out (Tensor, optional): the output tensor.

See also:

`torch.diagonal()` always returns the diagonal of its input.

`torch.diagflat()` always constructs a tensor with diagonal elements specified by the input.

Examples:

Get the square matrix where the input vector is the diagonal:

```
>>> a = torch.randn(3)
>>> a
tensor([ 0.5950,-0.0872, 2.3298])
>>> torch.diag(a)
tensor([[ 0.5950, 0.0000, 0.0000],
        [ 0.0000,-0.0872, 0.0000],
        [ 0.0000, 0.0000, 2.3298]])
>>> torch.diag(a, 1)
tensor([[ 0.0000, 0.5950, 0.0000, 0.0000],
        [ 0.0000, 0.0000,-0.0872, 0.0000],
        [ 0.0000, 0.0000, 0.0000, 2.3298],
        [ 0.0000, 0.0000, 0.0000, 0.0000]])
```

Get the k-th diagonal of a given matrix:

```
>>> a = torch.randn(3, 3)
>>> a
tensor([[-0.4264, 0.0255,-0.1064],
        [ 0.8795,-0.2429, 0.1374],
        [ 0.1029,-0.6482,-1.6300]])
```

(continues on next page)

(continued from previous page)

```
>>> torch.diag(a, 0)
tensor([-0.4264,-0.2429,-1.6300])
>>> torch.diag(a, 1)
tensor([ 0.0255, 0.1374])
```

hyperlearn.base.**diagSum**(*X*, *Y*, *transpose_a=False*)

hyperlearn.base.**dtype**(*tensor*)

hyperlearn.base.**einsum**(*notation*, **args*, *tensor=False*)

hyperlearn.base.**eps**(*X*)

hyperlearn.base.**isArray**(*X*)

hyperlearn.base.**isDict**(*X*)

hyperlearn.base.**isIterable**(*X*)

hyperlearn.base.**isList**(*X*)

hyperlearn.base.**isTensor**(*X*)

hyperlearn.base.**ones**(**size*, *out=None*, *dtype=None*, *layout=torch.strided*, *device=None*, *requires_grad=False*) → Tensor

Returns a tensor filled with the scalar value *1*, with the shape defined by the variable argument `size`.

Args:

size (int...): a sequence of integers defining the shape of the output tensor. Can be a variable number of arguments or a collection like a list or tuple.

out (Tensor, optional): the output tensor. dtype (`torch.dtype`, optional): the desired data type of returned tensor.

Default: if `None`, uses a global default (see `torch.set_default_tensor_type()`).

layout (`torch.layout`, optional): the desired layout of returned Tensor. Default: `torch.strided`.

device (`torch.device`, optional): the desired device of returned tensor. Default: if `None`, uses the current device for the default tensor type (see `torch.set_default_tensor_type()`). `device` will be the CPU for CPU tensor types and the current CUDA device for CUDA tensor types.

requires_grad (bool, optional): If autograd should record operations on the returned tensor. Default: `False`.

Example:

```
>>> torch.ones(2, 3)
tensor([[ 1.,  1.,  1.],
        [ 1.,  1.,  1.]])

>>> torch.ones(5)
tensor([ 1.,  1.,  1.,  1.,  1.])
```

hyperlearn.base.**resolution**(*X*)

hyperlearn.base.**return_numpy**(*args*)

hyperlearn.base.**return_torch**(*args*)

hyperlearn.base.**rowSum**(*X*, *Y=None*, *transpose_a=False*)

hyperlearn.base.**squareSum**(*X*)

hyperlearn.base.**stack**(**args*)

hyperlearn.base.**torch_dot**()
matmul(input, other, out=None) -> Tensor

Matrix product of two tensors.

The behavior depends on the dimensionality of the tensors as follows:

- If both tensors are 1-dimensional, the dot product (scalar) is returned.
- If both arguments are 2-dimensional, the matrix-matrix product is returned.
- If the first argument is 1-dimensional and the second argument is 2-dimensional, a 1 is prepended to its dimension for the purpose of the matrix multiply. After the matrix multiply, the prepended dimension is removed.
- If the first argument is 2-dimensional and the second argument is 1-dimensional, the matrix-vector product is returned.
- If both arguments are at least 1-dimensional and at least one argument is N-dimensional (where N > 2), then a batched matrix multiply is returned. If the first argument is 1-dimensional, a 1 is prepended to its dimension for the purpose of the batched matrix multiply and removed after. If the second argument is 1-dimensional, a 1 is appended to its dimension for the purpose of the batched matrix multiple and removed after. The non-matrix (i.e. batch) dimensions are broadcasted (and thus must be broadcastable). For example, if `input` is a $(j \times 1 \times n \times m)$ tensor and `other` is a $(k \times m \times p)$ tensor, `out` will be an $(j \times k \times n \times p)$ tensor.

Note: The 1-dimensional dot product version of this function does not support an `out` parameter.

Arguments: input (Tensor): the first tensor to be multiplied other (Tensor): the second tensor to be multiplied out (Tensor, optional): the output tensor.

Example:

```
>>> # vector x vector
>>> tensor1 = torch.randn(3)
>>> tensor2 = torch.randn(3)
>>> torch.matmul(tensor1, tensor2).size()
torch.Size([])
>>> # matrix x vector
>>> tensor1 = torch.randn(3, 4)
>>> tensor2 = torch.randn(4)
>>> torch.matmul(tensor1, tensor2).size()
torch.Size([3])
>>> # batched matrix x broadcasted vector
>>> tensor1 = torch.randn(10, 3, 4)
>>> tensor2 = torch.randn(4)
>>> torch.matmul(tensor1, tensor2).size()
torch.Size([10, 3])
>>> # batched matrix x batched matrix
>>> tensor1 = torch.randn(10, 3, 4)
>>> tensor2 = torch.randn(10, 4, 5)
>>> torch.matmul(tensor1, tensor2).size()
```

(continues on next page)

(continued from previous page)

```
torch.Size([10, 3, 5])
>>> # batched matrix x broadcasted matrix
>>> tensor1 = torch.randn(10, 3, 4)
>>> tensor2 = torch.randn(4, 5)
>>> torch.matmul(tensor1, tensor2).size()
torch.Size([10, 3, 5])
```

1.1.1.3 hyperlearn.linalg module

hyperlearn.linalg.**cholesky**(*XTX*, *alpha=None*, *fast=True*)

Computes the Cholesky Decompsition of a Hermitian Matrix (Positive Symmetric Matrix) giving a Upper Triangular Matrix.

Cholesky Decomposition is used as the default solver in HyperLearn, as it is super fast and allows regularization. HyperLearn's implementation also handles rank deficient and ill-conditioned matrices perfectly with the help of the limiting behaivour of adding forced epsilon regularization.

If USE_GPU: Uses PyTorch's Cholesky. Speed is OK.

If CPU: Uses Numpy's Fortran C based Cholesky. If NUMBA is not installed, uses very fast LAPACK functions.

Alpha is added for regularization purposes. This prevents system rounding errors and promises better convergence rates.

hyperlearn.linalg.**invCholesky**(*X*, *fast=False*)

Computes the Inverse of a Hermitian Matrix (Positive Symmetric Matrix) after provided with Cholesky's Lower Triangular Matrix.

This is used in conjunction in solveCholesky, where the inverse of the covariance matrix is needed.

If USE_GPU: Uses PyTorch's Triangular Solve given identity matrix. Speed is OK.

If CPU: Uses very fast LAPACK algorithms for triangular system inverses.

Note that LAPACK's single precision (float32) solver (strtri) is much more unstable than double (float64). So, default strtri is OFF. However, speeds are reduced by 50%.

hyperlearn.linalg.**pinvCholesky**(*X*, *alpha=None*, *fast=False*)

Computes the approximate pseudoinverse of any matrix using Cholesky Decomposition This means X @ pinv(X) approx = eye(n).

Note that this is super fast, and will be used in HyperLearn as the default pseudoinverse solver for any matrix. Care is taken to make the algorithm converge, and this is done via forced epsilon regularization.

HyperLearn's implementation also handles rank deficient and ill-conditioned matrices perfectly with the help of the limiting behaivour of adding forced epsilon regularization.

If USE_GPU: Uses PyTorch's Cholesky. Speed is OK.

If CPU: Uses Numpy's Fortran C based Cholesky. If NUMBA is not installed, uses very fast LAPACK functions.

Alpha is added for regularization purposes. This prevents system rounding errors and promises better convergence rates.

hyperlearn.linalg.**cholSolve**(*A*, *b*, *alpha=None*)

[Added 20/10/2018] Faster than direct inverse solve. Finds coefficients in linear regression allowing A @ theta = b. Notice auto adds epsilon jitter if solver fails.

hyperlearn.linalg.**svd**(*X*, *fast=True*, *U_decision=False*, *transpose=True*)

[Edited 9/11/2018 –> Modern Big Data Algorithms p/n ratio check] Computes the Singular Value Decomposition of any matrix. So, X = U * S @ VT. Note will compute svd(X.T) if p > n. Should be 99% same result. This means this implementation's time complexity is O[min(np^2, n^2p)]

If USE_GPU: Uses PyTorch's SVD. PyTorch uses (for now) a NON divide-n-conquer algo. Submitted report to PyTorch: https://github.com/pytorch/pytorch/issues/11174

If CPU: Uses Numpy's Fortran C based SVD. If NUMBA is not installed, uses divide-n-conqeur LAPACK functions.

If Transpose: Will compute if possible svd(X.T) instead of svd(X) if p > n. Default setting is TRUE to maintain speed.

SVD_Flip is used for deterministic output. Does NOT follow Sklearn convention. This flips the signs of U and VT, using VT_based decision.

hyperlearn.linalg.**lu**(*X*, *L_only=False*, *U_only=False*)

[Edited 8/11/2018 Changed to LAPACK LU if L/U only wanted] Computes the LU Decomposition of any matrix with pivoting. Provides L only or U only if specified.

Much faster than Scipy if only U/L wanted, and more memory efficient, since data is altered inplace.

hyperlearn.linalg.**qr**(*X*, *Q_only=False*, *R_only=False*, *overwrite=False*)

[Edited 8/11/2018 Added Q, R only parameters. Faster than Numba] [Edited 9/11/2018 Made R only more memory efficient (no data copying)] Computes the reduced QR Decomposition of any matrix. Uses optimized NUMBA QR if avaliable else use's Scipy's version.

Provides Q or R only if specified, and is must faster + more memory efficient since data is changed inplace.

hyperlearn.linalg.**pinv**(*X*, *alpha=None*, *fast=True*)

Computes the pseudoinverse of any matrix. This means X @ pinv(X) = eye(n).

Optional alpha is used for regularization purposes.

If USE_GPU: Uses PyTorch's SVD. PyTorch uses (for now) a NON divide-n-conquer algo. Submitted report to PyTorch: https://github.com/pytorch/pytorch/issues/11174

If CPU: Uses Numpy's Fortran C based SVD. If NUMBA is not installed, uses divide-n-conqeur LAPACK functions.

Condition number is: float32 = 1e3 * eps * max(S) float64 = 1e6 * eps * max(S)

hyperlearn.linalg.**pinvh**(*XTX*, *alpha=None*, *fast=True*)

Computes the pseudoinverse of a Hermitian Matrix (Positive Symmetric Matrix) using Eigendecomposition.

Uses the fact that the matrix is special, and so time complexity is approximately reduced by 1/2 or more when compared to full SVD.

If USE_GPU: Uses PyTorch's EIGH. PyTorch uses (for now) a non divide-n-conquer algo.

If CPU: Uses Numpy's Fortran C based EIGH. If NUMBA is not installed, uses very fast divide-n-conqeur LAPACK functions. Note Scipy's EIGH as of now is NON divide-n-conquer. Submitted report to Scipy: https://github.com/scipy/scipy/issues/9212

Condition number is: float32 = 1e3 * eps * max(abs(S)) float64 = 1e6 * eps * max(abs(S))

Alpha is added for regularization purposes. This prevents system rounding errors and promises better convergence rates.

hyperlearn.linalg.**eigh**(*XTX*, *alpha=None*, *fast=True*, *svd=False*, *positive=False*, *qr=False*)

Computes the Eigendecomposition of a Hermitian Matrix (Positive Symmetric Matrix).

Note: Slips eigenvalues / eigenvectors with MAX first. Scipy convention is MIN first, but MAX first is SVD convention.

Uses the fact that the matrix is special, and so time complexity is approximately reduced by 1/2 or more when compared to full SVD.

If POSITIVE is True, then all negative eigenvalues will be set to zero, and return value will be VT and not V.

If SVD is True, then eigenvalues will be square rooted as well.

If USE_GPU: Uses PyTorch's EIGH. PyTorch uses (for now) a non divide-n-conquer algo.

If CPU: Uses Numpy's Fortran C based EIGH. If NUMBA is not installed, uses very fast divide-n-conqeur LAPACK functions. Note Scipy's EIGH as of now is NON divide-n-conquer. Submitted report to Scipy: https://github.com/scipy/scipy/issues/9212

Alpha is added for regularization purposes. This prevents system rounding errors and promises better convergence rates.

Also uses eig_flip to flip the signs of the eigenvectors to ensure deterministic output.

hyperlearn.linalg.**pinvEig**(*X*, *alpha=None*, *fast=True*)

Computes the approximate pseudoinverse of any matrix X using Eigendecomposition on the covariance matrix XTX or XXT

Uses a special trick where: If n >= p: X^-1 approx = (XT @ X)^-1 @ XT If n < p: X^-1 approx = XT @ (X @ XT)^-1

If USE_GPU: Uses PyTorch's EIGH. PyTorch uses (for now) a non divide-n-conquer algo.

If CPU: Uses Numpy's Fortran C based EIGH. If NUMBA is not installed, uses very fast divide-n-conqeur LAPACK functions. Note Scipy's EIGH as of now is NON divide-n-conquer. Submitted report to Scipy: https://github.com/scipy/scipy/issues/9212

Condition number is: float32 = 1e3 * eps * max(abs(S)) float64 = 1e6 * eps * max(abs(S))

Alpha is added for regularization purposes. This prevents system rounding errors and promises better convergence rates.

hyperlearn.linalg.**eig**(*X*, *alpha=None*, *fast=True*, *U_decision=False*, *svd=False*, *stable=False*)

[Edited 8/11/2018 Made QR-SVD even faster –> changed to n >= p from n >= 5/3p] Computes the Eigendecomposition of any matrix using either QR then SVD or just SVD. This produces much more stable solutions that pure eigh(covariance), and thus will be necessary in some cases.

If STABLE is True, then EIGH will be bypassed, and instead SVD or QR/SVD will be used instead. This is to guarantee stability, since EIGH uses epsilon jitter along the diagonal of the covariance matrix.

If n >= 5/3 * p: Uses QR followed by SVD noticing that U is not needed. This means Q @ U is not required, reducing work.

Note Sklearn's Incremental PCA was used for the constant 5/3 [*Matrix Computations, Third Edition, G. Holub and C. Van Loan, Chapter 5, section 5.4.4, pp 252-253.*]

Else If n >= p: SVD is used, as QR would be slower.

Else If n <= p: SVD Transpose is used svd(X.T)

If stable is False: Eigh is used or SVD depending on the memory requirement.

Eig is the most stable Eigendecomposition in HyperLearn. It surpasses the stability of Eigh, as no epsilon jitter is added, unless specified when stable = False.

1.1.1.4 hyperlearn.utils module

class hyperlearn.utils.**lapack**(*function*, *fast=True*, *numba=None*)
Bases: object

[Added 11/11/2018] [Edited 13/11/2018 -> made into a class] [Edited 14/11/2018 -> fixed class] Get a LAPACK function based on the dtype(X). Acts like Scipy.

hyperlearn.utils.**svd_flip**(*U*, *VT*, *U_decision=True*)
Flips the signs of U and VT for SVD in order to force deterministic output.

Follows Sklearn convention by looking at U's maximum in columns as default.

hyperlearn.utils.**eig_flip**(*V*)
Flips the signs of V for Eigendecomposition in order to force deterministic output.

Follows Sklearn convention by looking at V's maximum in columns as default. This essentially mirrors svd_flip(U_decision = False)

hyperlearn.utils.**memoryXTX**(*X*)
Computes the memory usage for X.T @ X so that error messages can be broadcast without submitting to a memory error.

hyperlearn.utils.**memoryCovariance**(*X*)
Computes the memory usage for X.T @ X or X @ X.T so that error messages can be broadcast without submitting to a memory error.

hyperlearn.utils.**memorySVD**(*X*)
Computes the approximate memory usage of SVD(X) [transpose or not]. How it's computed:

X = U * S * VT U(n,p) * S(p) * VT(p,p) This means RAM usgae is np+p+p^2 approximately.

TODO: Divide N Conquer SVD vs old SVD

hyperlearn.utils.**traceXTX**
[Edited 18/10/2018] One drawback of truncated algorithms is that they can't output the correct variance explained ratios, since the full eigenvalue decomp needs to be done. However, using linear algebra, trace(XT*X) = sum(eigenvalues).

So, this function outputs the trace(XT*X) efficiently without computing explicitly XT*X.

Changes –> now uses Numba which is approx 20% faster.

hyperlearn.utils.**fastDot**(*A*, *B*, *C*)
[Added 23/9/2018] [Updated 1/10/2018 Error in calculating which is faster] Computes a fast matrix multiplication of 3 matrices. Either performs (A @ B) @ C or A @ (B @ C) depending which is more efficient.

hyperlearn.utils.**rowSum**(*X*, *norm=False*)
[Added 22/10/2018] Combines rowSum for matrices and arrays.

hyperlearn.utils.**rowSum_A**
[Added 22/10/2018] Computes rowSum**2 for dense array efficiently, instead of using einsum

hyperlearn.utils.**reflect**(*X*, *n_jobs=1*)
[Added 15/10/2018] [Edited 18/10/2018] Reflects lower triangular of matrix efficiently to upper. Notice much faster than say X += X.T or naive:

for i in range(n):

for j in range(i, n): X[i,j] = X[j,i]

In fact, it is much faster to perform vertically:

for i in range(1, n): Xi = X[i] for j in range(i):

X[j,i] = Xi[j]

The trick is to notice X[i], which reduces array access.

hyperlearn.utils.**addDiagonal**(*X*, *c=1*)
[Added 11/11/2018] Add c to diagonal of matrix

hyperlearn.utils.**setDiagonal**(*X*, *c=1*)
[Added 11/11/2018] Set c to diagonal of matrix

1.1.1.5 hyperlearn.random module

hyperlearn.random.**uniform**(*left*, *right*, *n*, *p=None*, *dtype=<class 'numpy.float32'>*)
[Added 6/11/2018]

Produces pseudo-random uniform numbers between left and right range. Notice much more memory efficient than Numpy, as provides a DTYPE argument (float32 supported).

hyperlearn.random.**uniform_vector**

1.1.1.6 hyperlearn.exceptions module

exception hyperlearn.exceptions.**FutureExceedsMemory**(*text='Operation done in the future uses more memory than what is free. HyperLearn'*)

Bases: BaseException

exception hyperlearn.exceptions.**PartialWrongShape**(*text='Partial SVD or Eig needs the same number of columns in both the previous iteration and the future iteration. Currenlty, the number of columns is different.'*)

Bases: BaseException

1.1.1.7 hyperlearn.multiprocessing module

1.1.1.8 hyperlearn.numba module

hyperlearn.numba.**svd**

hyperlearn.numba.**pinv**

hyperlearn.numba.**eigh**

hyperlearn.numba.**cholesky**

hyperlearn.numba.**lstsq**

hyperlearn.numba.**qr**

hyperlearn.numba.**norm**

hyperlearn.numba.**mean**(*X*, *axis=0*)

hyperlearn.numba.**sign**

hyperlearn.numba.**arange**

hyperlearn.numba.**minimum**

hyperlearn.numba.**maximum**

hyperlearn.numba.**multsum**

hyperlearn.numba.**squaresum**

1.1.1.9 hyperlearn.solvers module

hyperlearn.solvers.**solve**(*X*, *y*, *tol=1e-06*, *condition_limit=100000000.0*, *alpha=None*, *weights=None*, *copy=False*, *non_negative=False*, *max_iter=None*)

[As of 12/9/2018, an optional non_negative argument is added. Note as accurate as Scipy's NNLS, by copies ideas from gradient descent.]

[NOTE: as of 12/9/2018, LSMR is default in HyperLearn, replacing the 2nd fastest Cholesky Solve. LSMR is 2-4 times faster, and uses N less memory]

```
>>> WEIGHTS is an array of Weights for Weighted / Generalized Least Squares␣
↪[default None]
>>> theta = (XT*W*X)^-1*(XT*W*y)
```

Implements extremely fast least squares LSMR using orthogonalization as seen in Scipy's LSMR and https://arxiv.org/abs/1006.0758 [LSMR: An iterative algorithm for sparse least-squares problems] by David Fong, Michael Saunders.

Scipy's version of LSMR is surprisingly slow, as some slow design factors were used (ie np.sqrt(1 number) is slower than number**0.5, or min(a,b) is slower than using 1 if statement.)

ALPHA is provided for regularization purposes like Ridge Regression.

This algorithm works well for Sparse Matrices as well, and the time complexity analysis is approx: X.T @ y * min(n,p) times + 3 or so O(n) operations ==> O(np)*min(n,p) ==> either min(O(n^2p + n), O(np^2 + n)) **** Note if Weights is present, complexity increases.

Instead of fitting X^T*W*X, fits X^T/sqrt(W) So, O(np+n) is needed extra.

This complexity is much better than Cholesky Solve which is the next fastest in HyperLearn. Cholesky requires O(np^2) for XT * X, then Cholesky needs an extra 1/3*O(np^2), then inversion takes another 1/3*(np^2), and finally (XT*y) needs O(np).

So Cholesky needs O(5/3np^2 + np) >> min(O(n^2p + n), O(np^2 + n))

So by factor analysis, expect LSMR to be approx 2 times faster or so.

Interestingly, the Space Complexity is even more staggering. LSMR takes only maximum O(np^2) space for the computation of XT * y + some overhead.

**** Note if Weights is present, and COPY IS TRUE, then memory is DOUBLED.** Hence, try setting COPY to FALSE, memory will not change, and X will return back to its original state afterwards.

Cholesky requires XT * X space, which is already max O(n^2p) [which is huge]. Essentially, Cholesky shines when P is large, but N is small. LSMR is good for large N, medium P

Theta_hat = (XT * W * X)^-1 * (XT * y) In other words in gradient descent / iterative solves solve:

X * sqrt(W) * theta_hat = y * sqrt(W)

or: X*sqrt(W) ==> y*sqrt(W)

hyperlearn.solvers.**solveCholesky**(*X*, *y*, *alpha=None*, *fast=True*)

[Added 23/9/2018 added matrix multiplication decisions (faster multiply) ie: if (XTX)^1(XTy) or ((XTX)^-1XT)y is faster]

[Edited 20/10/2018 Major update - added LAPACK cholSolve –> 20% faster] [Edited 30/10/2018 Reduced RAM usage by clearing unused variables]

Computes the Least Squares solution to X @ theta = y using Cholesky Decomposition. This is the default solver in HyperLearn.

Cholesky Solving is used as the 2nd default solver [as of 12/9/2018, default has been switched to LSMR (called solve)] in HyperLearn, as it is super fast and allows regularization. HyperLearn's implementation also handles rank deficient and ill-conditioned matrices perfectly with the help of the limiting behaivour of adding forced epsilon regularization.

Optional alpha is used for regularization purposes.

Method | Operations | Factor * np^2 |

|————|————————|————————| | Cholesky | 1/3 * np^2 | 1/3 | | QR | p^3/3 + np^2 | 1 - p/3n | | SVD | p^3 + np^2 | 1 - p/n |

If USE_GPU: Uses PyTorch's Cholesky and Triangular Solve given identity matrix. Speed is OK.

If CPU: Uses Numpy's Fortran C based Cholesky. If NUMBA is not installed, uses very fast LAPACK functions. Also, uses very fast LAPACK algorithms for triangular system inverses.

Note that LAPACK's single precision (float32) solver (strtri) is much more unstable than double (float64). You might see stability problems if FAST = TRUE. Set it to FALSE if theres issues.

hyperlearn.solvers.**solveSVD**(*X*, *y*, *n_components=None*, *alpha=None*, *fast=True*)

[Edited 6/11/2018 Added n_components for Partial Solving] Computes the Least Squares solution to X @ theta = y using SVD. Slow, but most accurate. Specify n_components to reduce overfitting. Heurestic is 95% of variance is captured, if set to 'auto'.

Optional alpha is used for regularization purposes.

If USE_GPU: Uses PyTorch's SVD. PyTorch uses (for now) a NON divide-n-conquer algo. Submitted report to PyTorch: https://github.com/pytorch/pytorch/issues/11174

If CPU: Uses Numpy's Fortran C based SVD. If NUMBA is not installed, uses divide-n-conqeur LAPACK functions.

Condition number is: float32 = 1e3 * eps * max(S) float64 = 1e6 * eps * max(S)

hyperlearn.solvers.**solveEig**(*X*, *y*, *alpha=None*, *fast=True*)

[Edited 30/10/2018 Reduced RAM usage by clearing unused variables]

Computes the Least Squares solution to X @ theta = y using Eigendecomposition on the covariance matrix XTX or XXT. Medium speed and accurate, where this lies between SVD and Cholesky.

Optional alpha is used for regularization purposes.

If USE_GPU: Uses PyTorch's EIGH. PyTorch uses (for now) a non divide-n-conquer algo. Submitted report to PyTorch: https://github.com/pytorch/pytorch/issues/11174

If CPU: Uses Numpy's Fortran C based EIGH. If NUMBA is not installed, uses divide-n-conqeur LAPACK functions. Note Scipy's EIGH as of now is NON divide-n-conquer. Submitted report to Scipy: https://github.com/scipy/scipy/issues/9212

Condition number is: float32 = 1e3 * eps * max(abs(S)) float64 = 1e6 * eps * max(abs(S))

Alpha is added for regularization purposes. This prevents system rounding errors and promises better convergence rates.

hyperlearn.solvers.**solvePartial**(*X*, *y*, *n_components=None*, *alpha=None*, *fast=True*)
[Added 6/11/2018] Computes the Least Squares solution to X @ theta = y using Randomized SVD. Much faster than normal SVD solving, and is not prone is overfitting.

Optional alpha is used for regularization purposes.

hyperlearn.solvers.**lstsq**(*X*, *y*)
Returns normal Least Squares solution using LAPACK and Numba if installed. PyTorch will default to Cholesky Solve.

hyperlearn.solvers.**solveTLS**(*X*, *y*, *solver='truncated'*)
[Added 6/11/2018] Performs Total Least Squares based on the implementation in Wikipedia: https://en.wikipedia.org/wiki/Total_least_squares. The naming is rather deceptive, as it doesn't mean it'll yield better results than pure SVD solving. Normal linear regression assumes Y|X has gaussian noise. TLS assumes this AND X|Y has noise.

Two solvers - full, truncated. Truncated is much much faster, as smallest eigen component is needed. Full solver uses Eigendecomposition, which is much much slower, but more accurate.

1.1.1.10 hyperlearn.stats module

hyperlearn.stats.**corr**(*X*, *y*)

hyperlearn.stats.**qr_stats**(*Q*, *R*)
XTX^-1 = RT * R

h = diag Q * QT

mean(h) used for normalized leverage

hyperlearn.stats.**svd_stats**(*U*, *S*, *VT*)

$$\mathbf{XTX^{-1} = V \frac{1}{S^2} VT}$$

h = diag U * UT

mean(h) used for normalized leverage

hyperlearn.stats.**ridge_stats**(*U*, *S*, *VT*, *alpha=1*)

$$\mathbf{exp\_theta\_hat = diag\ V \frac{S^2}{S^2 + aI} VT}$$

$$\mathbf{var\_theta\_hat = diag\ V \frac{S^2}{(S^2 + aI)^2} VT}$$

$$\mathbf{XTX^{-1} = V \frac{1}{S^2 + aI} VT}$$

h = diag U ——— UT S^2 + aI

mean(h) used for normalized leverage

1.1.1.11 hyperlearn.big_data.base module

1.1.1.12 hyperlearn.big_data.incremental module

hyperlearn.big_data.incremental.**partialEig**(*batch*, *S2*, *V*, *ratio=1*, *solver='full'*, *tol=None*, *max_iter='auto'*)

Fits a partial Eigendecomp after given old eigenvalues S2 and old eigenvector components V.

Note that V will be used as the number of old components, so when calling truncated or randomized, will output a specific number of eigenvectors and eigenvalues.

Checks if new batch's size matches that of the old V.

Note that PartialEig has different solvers. Either choose:

1. **full** Solves full Eigendecompsition on the data. This is the most stable and will guarantee the most robust results. You can select the number of components to keep within the model later.
2. **truncated** This keeps the top K right eigenvectors and top k eigenvalues, as determined by n_components. Note full Eig is not called for the truncated case, but rather ARPACK is called.
3. **randomized** Same as truncated, but instead of using ARPACK, uses randomized Eig.

hyperlearn.big_data.incremental.**partialSVD**(*batch*, *S*, *VT*, *ratio=1*, *solver='full'*, *tol=None*, *max_iter='auto'*)

Fits a partial SVD after given old singular values S and old components VT.

Note that VT will be used as the number of old components, so when calling truncated or randomized, will output a specific number of eigenvectors and singular values.

Checks if new batch's size matches that of the old VT.

Note that PartialSVD has different solvers. Either choose:

1. **full** Solves full SVD on the data. This is the most stable and will guarantee the most robust results. You can select the number of components to keep within the model later.
2. **truncated** This keeps the top K right eigenvectors and top k right singular values, as determined by n_components. Note full SVD is not called for the truncated case, but rather ARPACK is called.
3. **randomized** Same as truncated, but instead of using ARPACK, uses randomized SVD.

Notice how Batch = U @ S @ VT. However, partialSVD returns S, VT, and not U. In order to get U, you might consider using the relation that X = U @ S @ VT, and approximating U by:

X = U @ S @ VT X @ V = U @ S (X @ V)/S = U

So, U = (X @ V)/S, so you can output U from (X @ V)/S

You can also get U partially and slowly using reverseU.

1.1.1.13 hyperlearn.big_data.lsmr module

hyperlearn.big_data.lsmr.**Orthogonalize**(*a*, *b*)

hyperlearn.big_data.lsmr.**floatType**(*dtype*)

hyperlearn.big_data.lsmr.**lsmr**(*X*, *y*, *tol=1e-06*, *condition_limit=100000000.0*, *alpha=0*, *threshold=1000000000000.0*, *non_negative=False*, *max_iter=None*)

[As of 12/9/2018, an optional non_negative argument is added. Note as accurate as Scipy's NNLS, by copies ideas from gradient descent.]

Implements extremely fast least squares LSMR using orthogonalization as seen in Scipy's LSMR and https://arxiv.org/abs/1006.0758 [LSMR: An iterative algorithm for sparse least-squares problems] by David Fong, Michael Saunders.

Scipy's version of LSMR is surprisingly slow, as some slow design factors were used (ie np.sqrt(1 number) is slower than number**0.5, or min(a,b) is slower than using 1 if statement.)

ALPHA is provided for regularization purposes like Ridge Regression.

This algorithm works well for Sparse Matrices as well, and the time complexity analysis is approx: X.T @ y * min(n,p) times + 3 or so O(n) operations ==> O(np)*min(n,p) ==> either min(O(n^2p + n), O(np^2 + n))

This complexity is much better than Cholesky Solve which is the next fastest in HyperLearn. Cholesky requires O(np^2) for XT * X, then Cholesky needs an extra 1/3*O(np^2), then inversion takes another 1/3*(np^2), and finally (XT*y) needs O(np).

So Cholesky needs O(5/3np^2 + np) >> min(O(n^2p + n), O(np^2 + n))

So by factor analysis, expect LSMR to be approx 2 times faster or so. Interestingly, the Space Complexity is even more staggering. LSMR takes only maximum O(np^2) space for the computation of XT * y + some overhead.

Cholesky requires XT * X space, which is already max O(n^2p) [which is huge]. Essentially, Cholesky shines when P is large, but N is small. LSMR is good for large N, medium P

1.1.1.14 hyperlearn.big_data.randomized module

hyperlearn.big_data.randomized.**randomizedEig**(*X*, *n_components=2*, *max_iter='auto'*, *solver='lu'*, *n_oversamples=10*)

[Edited 9/11/2018 Fixed Eig_Flip] HyperLearn's Randomized Eigendecomposition is an extension of Sklearn's randomized SVD. HyperLearn notices that the computation of U is not necessary, hence will use QR followed by SVD or just SVD depending on the situation.

Likewise, solver = LU is default, and follows randomizedSVD

References

- Sklearn's RandomizedSVD
- Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions Halko, et al., 2009 http://arxiv.org/abs/arXiv:0909.4061
- A randomized algorithm for the decomposition of matrices Per-Gunnar Martinsson, Vladimir Rokhlin and Mark Tygert
- An implementation of a randomized algorithm for principal component analysis A. Szlam et al. 2014

hyperlearn.big_data.randomized.**randomizedPinv**(*X*, *n_components=None*, *alpha=None*)

[Added 6/11/2018] Implements fast randomized pseudoinverse with regularization. Can be used as an approximation to the matrix inverse.

hyperlearn.big_data.randomized.**randomizedSVD**(*X*, *n_components=2*, *max_iter='auto'*, *solver='lu'*, *n_oversamples=10*)

[Edited 9/11/2018 Fixed SVD_flip] HyperLearn's Fast Randomized SVD is approx 10 - 30 % faster than Sklearn's implementation depending on n_components and max_iter.

Uses NUMBA Jit accelerated functions when available, and tries to reduce memory overhead by chaining operations.

Uses QR, LU or no solver to find the best SVD decomp. QR is most stable, but can be 2x slower than LU.

******n_oversamples = 10. This follows Sklearn convention to increase the chance** of more accurate SVD.

References

- Sklearn's RandomizedSVD
- Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions Halko, et al., 2009 http://arxiv.org/abs/arXiv:0909.4061
- A randomized algorithm for the decomposition of matrices Per-Gunnar Martinsson, Vladimir Rokhlin and Mark Tygert
- An implementation of a randomized algorithm for principal component analysis A. Szlam et al. 2014

hyperlearn.big_data.randomized.**randomized_projection**(*X*, *k*, *solver='lu'*, *max_iter=4*)

[Edited 8/11/2018 Added QR Q_only parameter] Projects X onto some random eigenvectors, then using a special variant of Orthogonal Iteration, finds the closest orthogonal representation for X.

Solver can be QR or LU or None.

1.1.1.15 hyperlearn.big_data.truncated module

hyperlearn.big_data.truncated.**truncatedEigh**(*XTX*, *n_components=2*, *tol=None*, *svd=False*, *which='largest'*)

[Edited 6/11/2018 Added smallest / largest command] Computes the Truncated Eigendecomposition of a Hermitian Matrix (positive definite). K = 2 for default. Return format is LARGEST eigenvalue first.

If SVD is True, then outputs S2**0.5 and sets negative S2 to 0 and outputs VT and not V.

Uses ARPACK from Scipy to compute the truncated decomp. Note that to make it slightly more stable and faster, follows Sklearn's random intialization from -1 -> 1.

Also note tolerance is resolution(X), and NOT eps(X)

Might switch to SPECTRA in the future.

EIGH FLIP is called to flip the eigenvector signs for deterministic output.

hyperlearn.big_data.truncated.**truncatedSVD**(*X*, *n_components=2*, *tol=None*, *transpose=True*, *U_decision=False*, *which='largest'*)

[Edited 6/11/2018 Added which command - can get largest or smallest eigen components] Computes the Truncated SVD of any matrix. K = 2 for default. Return format is LARGEST singular first first.

Uses ARPACK from Scipy to compute the truncated decomp. Note that to make it slightly more stable and faster, follows Sklearn's random intialization from -1 -> 1.

Also note tolerance is resolution(X), and NOT eps(X). Also note TRANSPOSE is True. This means instead of computing svd(X) if p > n, then computing svd(X.T) is faster, but you must output VT.T, S, U.T

Might switch to SPECTRA in the future.

SVD FLIP is called to flip the VT signs for deterministic output. Note uses VT based decision and not U based decision. U_decision can be changed to TRUE for Sklearn convention

hyperlearn.big_data.truncated.**truncatedEig**(*X*, *n_components=2*, *tol=None*, *svd=False*, *which='largest'*)

[Added 6/11/2018] Computes truncated eigendecomposition given any matrix X. Directly uses TruncatedSVD if memory is not enough, and returns eigen vectors/values. Also argument for smallest eigen components are provided.

1.1.1.16 hyperlearn.decomposition.base module

1.1.1.17 hyperlearn.decomposition.NMF module

1.1.1.18 hyperlearn.decomposition.PCA module

1.1.1.19 hyperlearn.decomposition.PCA module

1.1.1.20 hyperlearn.discriminant_analysis.base module

1.1.1.21 hyperlearn.discriminant_analysis.QDA module

1.1.1.22 hyperlearn.impute.SVDImpute module

hyperlearn.impute.SVDImpute.**fit**(*X*, *n_components='auto'*, *standardise=True*, *copy=True*)

[Added 31/10/2018] [Edited 2/11/2018 Fixed SVDImpute]

Fits a SVD onto the training data by projecting it to a lower space after being intially filled with column means. By default, n_components is determined automatically using log(p+1). Setting too low or too high mirrors mean imputation, and deletes the purpose of SVD imputation.

Returns: 1. S singular values 2. VT eigenvectors + mean, std, mins

hyperlearn.impute.SVDImpute.**transform**(*X*, *S*, *VT*, *mean*, *std*, *mins*, *standardise*, *copy=True*)

[Added 31/10/2018] [Edited 2/11/2018 FIxed SVDImpute]

The fundamental advantage of HyperLearn's SVD imputation is that a .transform method is provided. I do not require seeing the whole matrix for imputation, and can calculate SVD incrementally via the Incremental Module.

1.1.1.23 hyperlearn.metrics.cosine module

hyperlearn.metrics.cosine.**cosine_dis**

[Added 22/10/2018] Performs XXT*-1 + 1 quickly on the lower triangular part.

hyperlearn.metrics.cosine.**cosine_dis_triangular**

[Added 22/10/2018] Performs XXT*-1 + 1 quickly on the TCSR.

hyperlearn.metrics.cosine.**cosine_distances**(*X*, *Y=None*, *triangular=False*, *n_jobs=1*, *copy=False*)

[Added 15/10/2018] [Edited 18/10/2018] [Edited 22/10/2018 Added Y option] Note: when using Y, speed improvement is approx 5-10% only from Sklearn.

Slightly faster than Sklearn's Cosine Distances implementation. If you set triangular to TRUE, the result is much much faster. (Approx 50% faster than Sklearn)

hyperlearn.metrics.cosine.**cosine_distances_sparse**(*val*, *colPointer*, *rowIndices*, *n*, *p*, *triangular=False*, *dense_output=True*, *n_jobs=1*, *copy=True*)

[Added 22/10/2018] Slightly faster than Sklearn's Cosine Distances implementation.

If dense_output is set to FALSE, then a TCSR Matrix (Triangular CSR Matrix) is provided and not a CSR matrix. This has the advantage of using only 1/2n^2 - n memory and not n^2 memory.

hyperlearn.metrics.cosine.**cosine_sim_triangular**(*N*, *D*)

[Added 21/10/2018] Quickly performs X / norm_rows / norm_rows.T on the TCSR matrix.

hyperlearn.metrics.cosine.**cosine_sim_triangular_parallel**

[Added 21/10/2018] Quickly performs X / norm_rows / norm_rows.T on the TCSR matrix.

hyperlearn.metrics.cosine.**cosine_sim_triangular_single**

[Added 21/10/2018] Quickly performs X / norm_rows / norm_rows.T on the TCSR matrix.

hyperlearn.metrics.cosine.**cosine_similarity**(*X*, *Y=None*, *triangular=False*, *n_jobs=1*, *copy=False*)

[Added 20/10/2018] [Edited 22/201/2018] [Edited 22/10/2018 Added Y option] Note: when using Y, speed improvement is approx 5% only from Sklearn.

Cosine similarity is approx the same speed as Sklearn, but uses approx 10% less memory. One clear advantage is if you set triangular to TRUE, then it's faster.

hyperlearn.metrics.cosine.**cosine_similarity_sparse**(*val*, *colPointer*, *rowIndices*, *n*, *p*, *triangular=False*, *dense_output=True*, *n_jobs=1*, *copy=True*)

[Added 20/10/2018] [Edited 21/10/2018] Slightly faster than Sklearn's Cosine Similarity implementation.

If dense_output is set to FALSE, then a TCSR Matrix (Triangular CSR Matrix) is provided and not a CSR matrix. This has the advantage of using only 1/2n^2 - n memory and not n^2 memory.

1.1.1.24 hyperlearn.metrics.euclidean module

hyperlearn.metrics.euclidean.**euclidean_distances**(*X*, *Y=None*, *triangular=False*, *squared=False*, *n_jobs=1*)

[Added 15/10/2018] [Edited 16/10/2018] [Edited 22/10/2018 Added Y option] Notice: parsing in Y will result in only 10% - 15% speed improvement, not 30%.

Much much faster than Sklearn's implementation. Approx not 30% faster. Probably even faster if using n_jobs = -1. Uses the idea that distance(X, X) is symmetric, and thus algorithm runs only on 1/2 triangular part.

Old complexity: X @ XT n^2p rowSum(X^2) np XXT*-2 n^2 XXT+X^2 2n^2 maximum(XXT,0) n^2

n^2p + 4n^2 + np

New complexity: sym X @ XT n^2p/2 rowSum(X^2) np sym XXT*-2 n^2/2 sym XXT+X^2 n^2 maximum(XXT,0) n^2/2

n^2p/2 + 2n^2 + np

So New complexity approx= 1/2(Old complexity)

hyperlearn.metrics.euclidean.**euclidean_distances_sparse**(*val*, *colPointer*, *rowIndices*, *n*, *p*, *triangular=False*, *dense_output=True*, *squared=False*, *n_jobs=1*)

[Added 15/10/2018] [Edited 21/10/2018] Much much faster than Sklearn's implementation. Approx not 60% faster. Probably even faster if using n_jobs = -1 (actually 73% faster). [n = 10,000 p = 1,000] Uses the idea that distance(X, X) is symmetric, and thus algorithm runs only on 1/2 triangular part. Also notice memory usage is now 60% better than Sklearn.

If dense_output is set to FALSE, then a TCSR Matrix (Triangular CSR Matrix) is provided and not a CSR matrix. This has the advantage of using only 1/2n^2 - n memory and not n^2 memory.

Old complexity: X @ XT n^2p rowSum(X^2) np XXT*-2 n^2 XXT+X^2 2n^2 maximum(XXT,0) n^2

n^2p + 4n^2 + np

New complexity: sym X @ XT n^2p/2 rowSum(X^2) np sym XXT*-2 n^2/2 sym XXT+X^2 n^2 maximum(XXT,0) n^2/2

n^2p/2 + 2n^2 + np

So New complexity approx= 1/2(Old complexity)

hyperlearn.metrics.euclidean.**euclidean_triangular**(*S*, *D*, *squared=False*)
[Added 21/10/2018] Quickly performs -2D + X^2 + X.T^2 on the TCSR matrix. Also applies maximum(D, 0) and then square roots distances if required.

hyperlearn.metrics.euclidean.**euclidean_triangular_parallel**
[Added 21/10/2018] Quickly performs -2D + X^2 + X.T^2 on the TCSR matrix. Also applies maximum(D, 0) and then square roots distances if required.

hyperlearn.metrics.euclidean.**euclidean_triangular_single**
[Added 21/10/2018] Quickly performs -2D + X^2 + X.T^2 on the TCSR matrix. Also applies maximum(D, 0) and then square roots distances if required.

hyperlearn.metrics.euclidean.**maximum0**
[Added 15/10/2018] [Edited 21/10/2018] Computes maxmimum(XXT, 0) faster. Much faster than Sklearn since uses the notion that distance(X, X) is symmetric.

Steps:

maximum(XXT, 0) Optimised. Instead of n^2 operations, does n(n-1)/2 operations.

hyperlearn.metrics.euclidean.**maximum0_parallel**
[Added 15/10/2018] [Edited 21/10/2018] Computes maxmimum(XXT, 0) faster. Much faster than Sklearn since uses the notion that distance(X, X) is symmetric.

Steps:

maximum(XXT, 0) Optimised. Instead of n^2 operations, does n(n-1)/2 operations.

hyperlearn.metrics.euclidean.**mult_minus2**
[Added 17/10/2018] Quickly multiplies XXT by -2. Uses notion that XXT is symmetric, hence only lower triangular is multiplied.

1.1.1.25 hyperlearn.metrics.pairwise module

1.1.1.26 hyperlearn.sparse.base module

hyperlearn.sparse.base.**CreateCSR**(*X*, *n_jobs=1*)
[Added 10/10/2018] [Edited 13/10/2018] Much much faster than Scipy. In fact, HyperLearn uses less memory, by noticing indices >= 0, hence unsigned ints are used.

Likewise, parallelisation is seen possible with Numba with n_jobs. Notice, an error message will be provided if 20% of the data is only zeros. It needs to be more than 20% zeros for CSR Matrix to shine.

hyperlearn.sparse.base.**create_csr**(*X*, *rowCount*, *nnz*, *temp*)
[Added 10/10/2018] [Edited 13/10/2018] Before used extra memory keeping a Boolean Matrix (np bytes) and a ColIndex pointer which used p memory. Now, removed (np + p) memory usage, meaning larger matrices can be handled.

Algorithm is 3 fold:

1. Create RowIndices
2. **For every row in data:**
    3. Store until a non 0 is seen.

Algorithm takes approx O(n + np) time, which is similar to Scipy's. The only difference is now, parallelisation is possible, which can cut the time to approx O(n + np/c) where c = no of threads

hyperlearn.sparse.base.**create_csr_cache**
[Added 10/10/2018] [Edited 13/10/2018] Before used extra memory keeping a Boolean Matrix (np bytes) and a ColIndex pointer which used p memory. Now, removed (np + p) memory usage, meaning larger matrices can be handled.

Algorithm is 3 fold:

1. Create RowIndices
2. **For every row in data:**
    3. Store until a non 0 is seen.

Algorithm takes approx O(n + np) time, which is similar to Scipy's. The only difference is now, parallelisation is possible, which can cut the time to approx O(n + np/c) where c = no of threads

hyperlearn.sparse.base.**create_csr_parallel**
[Added 10/10/2018] [Edited 13/10/2018] Before used extra memory keeping a Boolean Matrix (np bytes) and a ColIndex pointer which used p memory. Now, removed (np + p) memory usage, meaning larger matrices can be handled.

Algorithm is 3 fold:

1. Create RowIndices
2. **For every row in data:**
    3. Store until a non 0 is seen.

Algorithm takes approx O(n + np) time, which is similar to Scipy's. The only difference is now, parallelisation is possible, which can cut the time to approx O(n + np/c) where c = no of threads

hyperlearn.sparse.base.**determine_nnz**
Uses close to no memory at all when computing how many non zeros are in the matrix. Notice the difference with Scipy is HyperLearn does NOT use nonzero(). This reduces memory usage dramatically.

hyperlearn.sparse.base.**getDtype**(*p*, *size*, *uint=True*)
Computes the exact best possible data type for CSR Matrix creation.

1.1.1.27 hyperlearn.sparse.csr module

hyperlearn.sparse.csr.**XXT_sparse**(*val*, *colPointer*, *rowIndices*, *n*, *p*)
See _XXT_sparse documentation.

hyperlearn.sparse.csr.**add_0**

hyperlearn.sparse.csr.**add_1**

hyperlearn.sparse.csr.**add_A**

hyperlearn.sparse.csr.**diagonal**
[Added 10/10/2018] [Edited 13/10/2018] Extracts the diagonal elements of a CSR Matrix. Note only gets square diagonal (not off-diagonal). HyperLearn's algorithm is faster than Scipy's, as it uses binary search, whilst Scipy uses Linear Search.

d = min(n, p) HyperLearn = O(d log p) Scipy = O(dp)

hyperlearn.sparse.csr.**diagonal_add**(*val*, *colPointer*, *rowIndices*, *n*, *p*, *addon*, *copy=True*)
See _diagonal_add documentation.

hyperlearn.sparse.csr.**div_0**

hyperlearn.sparse.csr.**div_1**

hyperlearn.sparse.csr.**div_A**

hyperlearn.sparse.csr.**get_element**
[Added 14/10/2018] Get A[i,j] element. HyperLearn's algorithm is O(logp) complexity, which is much better than Scipy's O(p) complexity. HyperLearn uses binary search to find the element.

hyperlearn.sparse.csr.**matT_mat**
Added [14/10/2018] A.T @ X is found where X is a dense matrix. Mostly the same as Scipy, albeit slightly faster. The difference is now, HyperLearn is parallelized, which can reduce times by 1/2 or more.

hyperlearn.sparse.csr.**matT_mat_parallel**
Added [14/10/2018] A.T @ X is found where X is a dense matrix. Mostly the same as Scipy, albeit slightly faster. The difference is now, HyperLearn is parallelized, which can reduce times by 1/2 or more.

hyperlearn.sparse.csr.**matT_vec**
Added [13/10/2018] X.T @ y is found. Notice how instead of converting CSR to CSC matrix, a direct X.T @ y can be found. Same complexity as mat_vec(X, y). Also, HyperLearn is parallelized, allowing for O(np/c) complexity.

hyperlearn.sparse.csr.**matT_vec_parallel**
Added [13/10/2018] X.T @ y is found. Notice how instead of converting CSR to CSC matrix, a direct X.T @ y can be found. Same complexity as mat_vec(X, y). Also, HyperLearn is parallelized, allowing for O(np/c) complexity.

hyperlearn.sparse.csr.**mat_mat**
Added [14/10/2018] A @ X is found where X is a dense matrix. Mostly the same as Scipy, albeit slightly faster. The difference is now, HyperLearn is parallelized, which can reduce times by 1/2 or more.

hyperlearn.sparse.csr.**mat_mat_parallel**
Added [14/10/2018] A @ X is found where X is a dense matrix. Mostly the same as Scipy, albeit slightly faster. The difference is now, HyperLearn is parallelized, which can reduce times by 1/2 or more.

hyperlearn.sparse.csr.**mat_vec**
Added [13/10/2018] X @ y is found. Scipy & HyperLearn has similar speed. Notice, now HyperLearn can be parallelised! This reduces complexity to approx O(np/c) where c = no of threads / cores

hyperlearn.sparse.csr.**mat_vec_parallel**
Added [13/10/2018] X @ y is found. Scipy & HyperLearn has similar speed. Notice, now HyperLearn can be parallelised! This reduces complexity to approx O(np/c) where c = no of threads / cores

hyperlearn.sparse.csr.**max_0**

hyperlearn.sparse.csr.**max_1**

hyperlearn.sparse.csr.**max_A**

hyperlearn.sparse.csr.**mean_0**

hyperlearn.sparse.csr.**mean_1**

hyperlearn.sparse.csr.**mean_A**

hyperlearn.sparse.csr.**min_0**

hyperlearn.sparse.csr.**min_1**

hyperlearn.sparse.csr.**min_A**

hyperlearn.sparse.csr.**mult_0**

hyperlearn.sparse.csr.**mult_1**

hyperlearn.sparse.csr.**mult_A**

hyperlearn.sparse.csr.**rowSum**
[Added 17/10/2018] Computes rowSum**2 for sparse matrix efficiently, instead of using einsum

hyperlearn.sparse.csr.**sum_0**

hyperlearn.sparse.csr.**sum_1**

hyperlearn.sparse.csr.**sum_A**

1.1.1.28 hyperlearn.sparse.tcsr module

1.1.1.29 Module contents

1.2 hyperlearn.base

The base module contains several decorators and methods.

1.2.1 Example code:

Singular Value Decomposition

```
import hyperlearn as hl
U, S, VT = hl.linalg.svd(X)
```

Pseudoinverse of a matrix using Cholesky Decomposition

```
from hyperlearn.linalg import pinvc
inv = pinvc(X)

# check if pinv(X) * X = identity
check = inv.dot(X)
```

QR Decomposition wanting only Q matrix

```
from hyperlearn import linalg
Q = linalg.qr(X, Q_only = True)

# check if Q == Q from full QR
q, r = linalg.qr(X)
```

1.3 hyperlearn.linalg

The linalg module contains all mathematical methods, decompositions and mirrors both Numpy's linalg and Scipy's linalg modules. HyperLearn's modules are all optimized and I also showcase some novel new algorithms.

1.3.1 Matrix Decompositions

Cholesky Decomposition	cholesky(X, [alpha])	:math:‘X=UsigmaV^T	Symmetric Square
LU Decomposition	lu(X, [L_only, U_only, overwrite])	X = L @ U	Any Matrix
Singular Value Decomposition	svd(X, [U_decision, overwrite])	X = U * S @ V.T	Any Matrix
QR Decomposition	qr(X, [Q_only, R_only, overwrite])	X = Q @ R	Any Matrix

1.3.2 Eigenvalue Problems

Symmetric EigenDecomposition	eigh(X, [alpha, svd, overwrite])	X = V * L @ V^-1	Symmetric Square

1.3.3 Matrix Inversion

Cholesky Inverse	cho_inv(X, [turbo])	inv(X) @ X = I	Symmetric Square
Pseudoinverse (Cholesky)	pinvc(X, [alpha, turbo])	inv(X) @ X = I	Any Matrix
Symmetric Pseudoinverse	pinvh(X, [alpha, turbo, n_jobs])	inv(X) @ X = I	Symmetric Square
Pseudoinverse (SVD)	pinv(X, [alpha, overwrite])	inv(X) @ X = I	Any Matrix
Pseudoinverse (LU Decomp)	pinvl(X, [alpha, turbo, overwrite])	inv(X) @ X = I	Square Matrix

1.4 License

GNU GENERAL PUBLIC LICENSE Version 3, 29 June 2007 ———————————

1.4.1 Preamble

The GNU General Public License is a free, copyleft license for software and other kinds of works.

The licenses for most software and other practical works are designed to take away your freedom to share and change the works. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change all versions of a program–to make sure it remains free software for all its users. We, the Free Software Foundation, use the GNU General Public License for most of our software; it applies also to any other work released this way by its authors. You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for them if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs, and that you know you can do these things.

To protect your rights, we need to prevent others from denying you these rights or asking you to surrender the rights. Therefore, you have certain responsibilities if you distribute copies of the software, or if you modify it: responsibilities to respect the freedom of others.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must pass on to the recipients the same freedoms that you received. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

Developers that use the GNU GPL protect your rights with two steps: (1) assert copyright on the software, and (2) offer you this License giving you legal permission to copy, distribute and/or modify it.

For the developers' and authors' protection, the GPL clearly explains that there is no warranty for this free software. For both users' and authors' sake, the GPL requires that modified versions be marked as changed, so that their problems will not be attributed erroneously to authors of previous versions.

Some devices are designed to deny users access to install or run modified versions of the software inside them, although the manufacturer can do so. This is fundamentally incompatible with the aim of protecting users' freedom to change the software. The systematic pattern of such abuse occurs in the area of products for individuals to use, which is precisely where it is most unacceptable. Therefore, we have designed this version of the GPL to prohibit the practice for those products. If such problems arise substantially in other domains, we stand ready to extend this provision to those domains in future versions of the GPL, as needed to protect the freedom of users.

Finally, every program is threatened constantly by software patents. States should not allow patents to restrict development and use of software on general-purpose computers, but in those that do, we wish to avoid the special danger that patents applied to a free program could make it effectively proprietary. To prevent this, the GPL assures that patents cannot be used to render the program non-free.

The precise terms and conditions for copying, distribution and modification follow.

1.4.2 TERMS AND CONDITIONS

0. Definitions.

"This License" refers to version 3 of the GNU General Public License.

"Copyright" also means copyright-like laws that apply to other kinds of works, such as semiconductor masks.

"The Program" refers to any copyrightable work licensed under this License. Each licensee is addressed as "you". "Licensees" and "recipients" may be individuals or organizations.

To "modify" a work means to copy from or adapt all or part of the work in a fashion requiring copyright permission, other than the making of an exact copy. The resulting work is called a "modified version" of the earlier work or a work "based on" the earlier work.

A "covered work" means either the unmodified Program or a work based on the Program.

To "propagate" a work means to do anything with it that, without permission, would make you directly or secondarily liable for infringement under applicable copyright law, except executing it on a computer or modifying a private copy. Propagation includes copying, distribution (with or without modification), making available to the public, and in some countries other activities as well.

To "convey" a work means any kind of propagation that enables other parties to make or receive copies. Mere interaction with a user through a computer network, with no transfer of a copy, is not conveying.

An interactive user interface displays "Appropriate Legal Notices" to the extent that it includes a convenient and prominently visible feature that (1) displays an appropriate copyright notice, and (2) tells the user that there is no

warranty for the work (except to the extent that warranties are provided), that licensees may convey the work under this License, and how to view a copy of this License. If the interface presents a list of user commands or options, such as a menu, a prominent item in the list meets this criterion.

1. Source Code.

The "source code" for a work means the preferred form of the work for making modifications to it. "Object code" means any non-source form of a work.

A "Standard Interface" means an interface that either is an official standard defined by a recognized standards body, or, in the case of interfaces specified for a particular programming language, one that is widely used among developers working in that language.

The "System Libraries" of an executable work include anything, other than the work as a whole, that (a) is included in the normal form of packaging a Major Component, but which is not part of that Major Component, and (b) serves only to enable use of the work with that Major Component, or to implement a Standard Interface for which an implementation is available to the public in source code form. A "Major Component", in this context, means a major essential component (kernel, window system, and so on) of the specific operating system (if any) on which the executable work runs, or a compiler used to produce the work, or an object code interpreter used to run it.

The "Corresponding Source" for a work in object code form means all the source code needed to generate, install, and (for an executable work) run the object code and to modify the work, including scripts to control those activities. However, it does not include the work's System Libraries, or general-purpose tools or generally available free programs which are used unmodified in performing those activities but which are not part of the work. For example, Corresponding Source includes interface definition files associated with source files for the work, and the source code for shared libraries and dynamically linked subprograms that the work is specifically designed to require, such as by intimate data communication or control flow between those subprograms and other parts of the work.

The Corresponding Source need not include anything that users can regenerate automatically from other parts of the Corresponding Source.

The Corresponding Source for a work in source code form is that same work.

2. Basic Permissions.

All rights granted under this License are granted for the term of copyright on the Program, and are irrevocable provided the stated conditions are met. This License explicitly affirms your unlimited permission to run the unmodified Program. The output from running a covered work is covered by this License only if the output, given its content, constitutes a covered work. This License acknowledges your rights of fair use or other equivalent, as provided by copyright law.

You may make, run and propagate covered works that you do not convey, without conditions so long as your license otherwise remains in force. You may convey covered works to others for the sole purpose of having them make modifications exclusively for you, or provide you with facilities for running those works, provided that you comply with the terms of this License in conveying all material for which you do not control copyright. Those thus making or running the covered works for you must do so exclusively on your behalf, under your direction and control, on terms that prohibit them from making any copies of your copyrighted material outside their relationship with you.

Conveying under any other circumstances is permitted solely under the conditions stated below. Sublicensing is not allowed; section 10 makes it unnecessary.

3. Protecting Users' Legal Rights From Anti-Circumvention Law.

No covered work shall be deemed part of an effective technological measure under any applicable law fulfilling obligations under article 11 of the WIPO copyright treaty adopted on 20 December 1996, or similar laws prohibiting or restricting circumvention of such measures.

When you convey a covered work, you waive any legal power to forbid circumvention of technological measures to the extent such circumvention is effected by exercising rights under this License with respect to the covered work, and you disclaim any intention to limit operation or modification of the work as a means of enforcing, against the work's users, your or third parties' legal rights to forbid circumvention of technological measures.

4. Conveying Verbatim Copies.

You may convey verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice; keep intact all notices stating that this License and any non-permissive terms added in accord with section 7 apply to the code; keep intact all notices of the absence of any warranty; and give all recipients a copy of this License along with the Program.

You may charge any price or no price for each copy that you convey, and you may offer support or warranty protection for a fee.

5. Conveying Modified Source Versions.

You may convey a work based on the Program, or the modifications to produce it from the Program, in the form of source code under the terms of section 4, provided that you also meet all of these conditions:

a) The work must carry prominent notices stating that you modified it, and giving a relevant date.

b) The work must carry prominent notices stating that it is released under this License and any conditions added under section 7. This requirement modifies the requirement in section 4 to "keep intact all notices".

c) You must license the entire work, as a whole, under this License to anyone who comes into possession of a copy. This License will therefore apply, along with any applicable section 7 additional terms, to the whole of the work, and all its parts, regardless of how they are packaged. This License gives no permission to license the work in any other way, but it does not invalidate such permission if you have separately received it.

d) If the work has interactive user interfaces, each must display Appropriate Legal Notices; however, if the Program has interactive interfaces that do not display Appropriate Legal Notices, your work need not make them do so.

A compilation of a covered work with other separate and independent works, which are not by their nature extensions of the covered work, and which are not combined with it such as to form a larger program, in or on a volume of a storage or distribution medium, is called an "aggregate" if the compilation and its resulting copyright are not used to limit the access or legal rights of the compilation's users beyond what the individual works permit. Inclusion of a covered work in an aggregate does not cause this License to apply to the other parts of the aggregate.

6. Conveying Non-Source Forms.

You may convey a covered work in object code form under the terms of sections 4 and 5, provided that you also convey the machine-readable Corresponding Source under the terms of this License, in one of these ways:

a) Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by the Corresponding Source fixed on a durable physical medium customarily used for software interchange.

b) Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by a written offer, valid for at least three years and valid for as long as you offer spare parts or customer support for that product model, to give anyone who possesses the object code either (1) a copy of the Corresponding Source for all the software in the product that is covered by this License, on a durable physical medium customarily used for software interchange, for a price no more than your reasonable cost of physically performing this conveying of source, or (2) access to copy the Corresponding Source from a network server at no charge.

c) Convey individual copies of the object code with a copy of the written offer to provide the Corresponding Source. This alternative is allowed only occasionally and noncommercially, and only if you received the object code with such an offer, in accord with subsection 6b.

d) Convey the object code by offering access from a designated place (gratis or for a charge), and offer equivalent access to the Corresponding Source in the same way through the same place at no further charge. You need not require recipients to copy the Corresponding Source along with the object code. If the place to copy the object code is a network server, the Corresponding Source may be on a different server (operated by you or a third party) that supports equivalent copying facilities, provided you maintain clear directions next to the object code saying where to find the Corresponding Source. Regardless of what server hosts the Corresponding Source, you remain obligated to ensure that it is available for as long as needed to satisfy these requirements.

e) Convey the object code using peer-to-peer transmission, provided you inform other peers where the object code and Corresponding Source of the work are being offered to the general public at no charge under subsection 6d.

A separable portion of the object code, whose source code is excluded from the Corresponding Source as a System Library, need not be included in conveying the object code work.

A "User Product" is either (1) a "consumer product", which means any tangible personal property which is normally used for personal, family, or household purposes, or (2) anything designed or sold for incorporation into a dwelling. In determining whether a product is a consumer product, doubtful cases shall be resolved in favor of coverage. For a particular product received by a particular user, "normally used" refers to a typical or common use of that class of product, regardless of the status of the particular user or of the way in which the particular user actually uses, or expects or is expected to use, the product. A product is a consumer product regardless of whether the product has substantial commercial, industrial or non-consumer uses, unless such uses represent the only significant mode of use of the product.

"Installation Information" for a User Product means any methods, procedures, authorization keys, or other information required to install and execute modified versions of a covered work in that User Product from a modified version of its Corresponding Source. The information must suffice to ensure that the continued functioning of the modified object code is in no case prevented or interfered with solely because modification has been made.

If you convey an object code work under this section in, or with, or specifically for use in, a User Product, and the conveying occurs as part of a transaction in which the right of possession and use of the User Product is transferred to the recipient in perpetuity or for a fixed term (regardless of how the transaction is characterized), the Corresponding Source conveyed under this section must be accompanied by the Installation Information. But this requirement does not apply if neither you nor any third party retains the ability to install modified object code on the User Product (for example, the work has been installed in ROM).

The requirement to provide Installation Information does not include a requirement to continue to provide support service, warranty, or updates for a work that has been modified or installed by the recipient, or for the User Product in which it has been modified or installed. Access to a network may be denied when the modification itself materially and adversely affects the operation of the network or violates the rules and protocols for communication across the network.

Corresponding Source conveyed, and Installation Information provided, in accord with this section must be in a format that is publicly documented (and with an implementation available to the public in source code form), and must require no special password or key for unpacking, reading or copying.

7. Additional Terms.

"Additional permissions" are terms that supplement the terms of this License by making exceptions from one or more of its conditions. Additional permissions that are applicable to the entire Program shall be treated as though they were included in this License, to the extent that they are valid under applicable law. If additional permissions apply only to part of the Program, that part may be used separately under those permissions, but the entire Program remains governed by this License without regard to the additional permissions.

When you convey a copy of a covered work, you may at your option remove any additional permissions from that copy, or from any part of it. (Additional permissions may be written to require their own removal in certain cases when you modify the work.) You may place additional permissions on material, added by you to a covered work, for which you have or can give appropriate copyright permission.

Notwithstanding any other provision of this License, for material you add to a covered work, you may (if authorized by the copyright holders of that material) supplement the terms of this License with terms:

a) Disclaiming warranty or limiting liability differently from the terms of sections 15 and 16 of this License; or

b) Requiring preservation of specified reasonable legal notices or author attributions in that material or in the Appropriate Legal Notices displayed by works containing it; or

c) Prohibiting misrepresentation of the origin of that material, or requiring that modified versions of such material be marked in reasonable ways as different from the original version; or

d) Limiting the use for publicity purposes of names of licensors or authors of the material; or

e) Declining to grant rights under trademark law for use of some trade names, trademarks, or service marks; or

f) Requiring indemnification of licensors and authors of that material by anyone who conveys the material (or modified versions of it) with contractual assumptions of liability to the recipient, for any liability that these contractual assumptions directly impose on those licensors and authors.

All other non-permissive additional terms are considered "further restrictions" within the meaning of section 10. If the Program as you received it, or any part of it, contains a notice stating that it is governed by this License along with a term that is a further restriction, you may remove that term. If a license document contains a further restriction but permits relicensing or conveying under this License, you may add to a covered work material governed by the terms of that license document, provided that the further restriction does not survive such relicensing or conveying.

If you add terms to a covered work in accord with this section, you must place, in the relevant source files, a statement of the additional terms that apply to those files, or a notice indicating where to find the applicable terms.

Additional terms, permissive or non-permissive, may be stated in the form of a separately written license, or stated as exceptions; the above requirements apply either way.

8. Termination.

You may not propagate or modify a covered work except as expressly provided under this License. Any attempt otherwise to propagate or modify it is void, and will automatically terminate your rights under this License (including any patent licenses granted under the third paragraph of section 11).

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, you do not qualify to receive new licenses for the same material under section 10.

9. Acceptance Not Required for Having Copies.

You are not required to accept this License in order to receive or run a copy of the Program. Ancillary propagation of a covered work occurring solely as a consequence of using peer-to-peer transmission to receive a copy likewise does not require acceptance. However, nothing other than this License grants you permission to propagate or modify any covered work. These actions infringe copyright if you do not accept this License. Therefore, by modifying or propagating a covered work, you indicate your acceptance of this License to do so.

10. Automatic Licensing of Downstream Recipients.

Each time you convey a covered work, the recipient automatically receives a license from the original licensors, to run, modify and propagate that work, subject to this License. You are not responsible for enforcing compliance by third parties with this License.

An "entity transaction" is a transaction transferring control of an organization, or substantially all assets of one, or subdividing an organization, or merging organizations. If propagation of a covered work results from an entity transaction, each party to that transaction who receives a copy of the work also receives whatever licenses to the work the party's predecessor in interest had or could give under the previous paragraph, plus a right to possession of the Corresponding Source of the work from the predecessor in interest, if the predecessor has it or can get it with reasonable efforts.

You may not impose any further restrictions on the exercise of the rights granted or affirmed under this License. For example, you may not impose a license fee, royalty, or other charge for exercise of rights granted under this License, and you may not initiate litigation (including a cross-claim or counterclaim in a lawsuit) alleging that any patent claim is infringed by making, using, selling, offering for sale, or importing the Program or any portion of it.

11. Patents.

A "contributor" is a copyright holder who authorizes use under this License of the Program or a work on which the Program is based. The work thus licensed is called the contributor's "contributor version".

A contributor's "essential patent claims" are all patent claims owned or controlled by the contributor, whether already acquired or hereafter acquired, that would be infringed by some manner, permitted by this License, of making, using, or selling its contributor version, but do not include claims that would be infringed only as a consequence of further modification of the contributor version. For purposes of this definition, "control" includes the right to grant patent sublicenses in a manner consistent with the requirements of this License.

Each contributor grants you a non-exclusive, worldwide, royalty-free patent license under the contributor's essential patent claims, to make, use, sell, offer for sale, import and otherwise run, modify and propagate the contents of its contributor version.

In the following three paragraphs, a "patent license" is any express agreement or commitment, however denominated, not to enforce a patent (such as an express permission to practice a patent or covenant not to sue for patent infringement). To "grant" such a patent license to a party means to make such an agreement or commitment not to enforce a patent against the party.

If you convey a covered work, knowingly relying on a patent license, and the Corresponding Source of the work is not available for anyone to copy, free of charge and under the terms of this License, through a publicly available network server or other readily accessible means, then you must either (1) cause the Corresponding Source to be so available, or (2) arrange to deprive yourself of the benefit of the patent license for this particular work, or (3) arrange, in a manner consistent with the requirements of this License, to extend the patent license to downstream recipients. "Knowingly relying" means you have actual knowledge that, but for the patent license, your conveying the covered work in a country, or your recipient's use of the covered work in a country, would infringe one or more identifiable patents in that country that you have reason to believe are valid.

If, pursuant to or in connection with a single transaction or arrangement, you convey, or propagate by procuring conveyance of, a covered work, and grant a patent license to some of the parties receiving the covered work authorizing them to use, propagate, modify or convey a specific copy of the covered work, then the patent license you grant is automatically extended to all recipients of the covered work and works based on it.

A patent license is "discriminatory" if it does not include within the scope of its coverage, prohibits the exercise of, or is conditioned on the non-exercise of one or more of the rights that are specifically granted under this License. You may not convey a covered work if you are a party to an arrangement with a third party that is in the business of distributing software, under which you make payment to the third party based on the extent of your activity of conveying the work, and under which the third party grants, to any of the parties who would receive the covered work from you, a discriminatory patent license (a) in connection with copies of the covered work conveyed by you (or copies made from those copies), or (b) primarily for and in connection with specific products or compilations that contain the covered work, unless you entered into that arrangement, or that patent license was granted, prior to 28 March 2007.

Nothing in this License shall be construed as excluding or limiting any implied license or other defenses to infringement that may otherwise be available to you under applicable patent law.

12. No Surrender of Others' Freedom.

If conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot convey a covered work so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not convey it at all. For example, if you agree to terms that obligate you to collect a royalty for further conveying from those to whom you convey the Program, the only way you could satisfy both those terms and this License would be to refrain entirely from conveying the Program.

13. Use with the GNU Affero General Public License.

Notwithstanding any other provision of this License, you have permission to link or combine any covered work with a work licensed under version 3 of the GNU Affero General Public License into a single combined work, and to convey the resulting work. The terms of this License will continue to apply to the part which is the covered work, but the

special requirements of the GNU Affero General Public License, section 13, concerning interaction through a network will apply to the combination as such.

14. Revised Versions of this License.

The Free Software Foundation may publish revised and/or new versions of the GNU General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies that a certain numbered version of the GNU General Public License "or any later version" applies to it, you have the option of following the terms and conditions either of that numbered version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of the GNU General Public License, you may choose any version ever published by the Free Software Foundation.

If the Program specifies that a proxy can decide which future versions of the GNU General Public License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Program.

Later license versions may give you additional or different permissions. However, no additional obligations are imposed on any author or copyright holder as a result of your choosing to follow a later version.

15. Disclaimer of Warranty.

THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

16. Limitation of Liability.

IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MODIFIES AND/OR CONVEYS THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

17. Interpretation of Sections 15 and 16.

If the disclaimer of warranty and limitation of liability provided above cannot be given local legal effect according to their terms, reviewing courts shall apply local law that most closely approximates an absolute waiver of all civil liability in connection with the Program, unless a warranty or assumption of liability accompanies a copy of the Program in return for a fee.

END OF TERMS AND CONDITIONS

How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.

To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively state the exclusion of warranty; and each file should have at least the "copyright" line and a pointer to where the full notice is found.

<one line to give the program's name and a brief idea of what it does.> Copyright (C) <year> <name of author>

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program. If not, see <https://www.gnu.org/licenses/>.

Also add information on how to contact you by electronic and paper mail.

If the program does terminal interaction, make it output a short notice like this when it starts in an interactive mode:

<program> Copyright (C) <year> <name of author> This program comes with ABSOLUTELY NO WARRANTY; for details type ‘show w’. This is free software, and you are welcome to redistribute it under certain conditions; type ‘show c’ for details.

The hypothetical commands ‘show w’ and ‘show c’ should show the appropriate parts of the General Public License. Of course, your program’s commands might be different; for a GUI interface, you would use an “about box”.

You should also get your employer (if you work as a programmer) or school, if any, to sign a “copyright disclaimer” for the program, if necessary. For more information on this, and how to apply and follow the GNU GPL, see <https://www.gnu.org/licenses/>.

The GNU General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Lesser General Public License instead of this License. But first, please read <https://www.gnu.org/licenses/why-not-lgpl.html>.

1.5 Contact

Please email me for licensing questions.

CHAPTER 2

Directory

- genindex
- modindex
- search

Python Module Index

h

Index

H

I

L

M

N

O

P

Q

R

S

T

U

X