Greedy Algorithms for Linear Dimensionality Reduction

Sihui Wang

1. Introduction

We have implemented a greedy algorithm for linear dimensionality reduction with the objective of maximizing Fisher information distance. Now we will show that we could also adopt this framework to solve the problem of linear dimensionality reduction with the objective of maximizing KL divergence, symmetric divergence, and Hellinger distance.

The generic framework of greedy algorithms is:

Greedy Algorithms for Linear Dimensionality Reduction

Input: $\mu_1, \mu_2, \Sigma_1, \Sigma_2, r$

Output: $A \in \mathbb{R}^{r \times n}$

1. Initialization. Solution vectors: $U \leftarrow \emptyset$;

Transformer: $T \leftarrow \Sigma_1^{-\frac{1}{2}}$; $\mu \leftarrow T(\mu_2 - \mu_1)$; $\Sigma \leftarrow T\Sigma_2 T$;

2.For k=1 to r

(1) Initialization step:

<1.1> Generate a non-zero vector u;

<1.2> Project u into the subspace: $u_{k,\perp} \leftarrow u - \sum_{i=1}^{k-1} < u, u_i > u_i$;

<1.3> Normalization: $u_{k,0} \leftarrow \frac{u_{k,\perp}}{\|u_{k,\perp}\|_2}$;

<1.4> $l \leftarrow 0$;

(2) **Do**

<2.1> Compute the gradient in Euclidean space: calculate $\nabla_{u_{k,l}} F$;

<2.2> Compute the gradient on the manifold: $\nabla^M_{u_{k,l}} F \leftarrow \nabla_{u_{k,l}} F - < \nabla_{u_{k,l}} F, u_{k,l} > u_{k,l}$;

<2.3> Gradient projection: $\nabla^{\perp}_{u_{k,l}} F \leftarrow \nabla^M_{u_{k,l}} F - \sum_{i=1}^{k-1} < \nabla^M_{u_{k,l}} F, u_i > u_i$;

<2.4> Update $u_{k,l}$ in the usual way: $u^*_{k,l} \leftarrow u_{k,l} + \alpha \nabla^{\perp}_{u_{k,l}} F$;

<2.5> Retraction: $u_{k,l+1} = \frac{u^*_{k,l}}{\|u^*_{k,l}\|_2}$

<2.6> $l \leftarrow l + 1$;

While not Convergence;

(3) Obtain $u_k \leftarrow u_{k,l}$ after the convergence of $u_{k,l}$; $U \leftarrow U \cup \{u_k\}$;

End For

3. $A^* \leftarrow (u_1^T, \ldots, u_r^T)^T$;

Return $A \leftarrow A^* T$

where F could be D_{KL}, D_{SKL}, D_H, or D_F.

2. Motivation

Previously we have already had gradient-based algorithms which treat linear dimensionality reduction as optimization on Grassmann manifolds, so why do I want to find another heuristic algorithm?

One reason is efficiency. Suppose for some $X\sim\mathcal{N}(\mu_1,\Sigma_1)$, $Y\sim\mathcal{N}(\mu_2,\Sigma_2)$, we need a two dimensional solution after we obtained a one dimensional solution. With previous algorithms we need to start all over again, while with greedy algorithms the new two dimensional solutions are built upon the existing one dimensional solutions, which could potentially save some computation time.

Sometimes we might want to run linear dimensionality reduction for multiple times in order to decide the optimal and sufficient dimension for dimensionality reduction. In this case, it is obvious that the greedy algorithm is much more efficient than the previous algorithms.

Another reason is interpretability. With previous algorithms, it is difficult for us to find out, for example, that there is any connection between a one dimensional solution and a two dimensional solution for the same $X\sim\mathcal{N}(\mu_1,\Sigma_1)$, $Y\sim\mathcal{N}(\mu_2,\Sigma_2)$. Meanwhile, with greedy algorithms it is easy for us to decompose a solution into several components of one dimensional solutions that are orthogonal to each other, which is easier for us to explain and understand.

Although these greedy algorithms are heuristic algorithms, it is shown in the experiments that they don't sacrifice much in their performances, which is why I think that they are worth noting and their advantages outweigh their shortcomings.

3. Experiments

We could see from experiments that the overall performance of greedy algorithms is comparable to the performance of our previous algorithms.

Here I only provide a sample of the results. For full results please refer to the supplemental documents.

Abalone, Class 1 vs Class 2, D_{KL}				Abalone, Class 1 vs Class 2, D_{KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	1180/1528	213/1307	0.93522	$r=1$	1176/1528	212/1307	0.93184
$r=2$	288/1528	1000/1307	0.95359	$r=2$	288/1528	1001/1307	0.95436
$r=3$	320/1528	927/1307	0.91868	$r=3$	318/1528	925/1307	0.91584
Abalone, Class 1 vs Class 2, D_{r-KL}				Abalone, Class 1 vs Class 2, D_{r-KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	1176/1528	212/1307	0.93184	$r=1$	168/1528	1126/1307	0.97146
$r=2$	288/1528	1001/1307	0.95436	$r=2$	288/1528	1000/1307	0.95359
$r=3$	343/1528	941/1307	0.94445	$r=3$	319/1528	923/1307	0.91497
Abalone, Class 1 vs Class 2, D_{SKL}				Abalone, Class 1 vs Class 2, D_{SKL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	1176/1528	212/1307	0.93184	$r=1$	168/1528	1126/1307	0.97146
$r=2$	288/1528	1001/1307	0.95436	$r=2$	288/1528	1001/1307	0.95436
$r=3$	320/1528	925/1307	0.91715	$r=3$	318/1528	926/1307	0.91661
Abalone, Class 1 vs Class 2, D_H				Abalone, Class 1 vs Class 2, D_H, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	1178/1528	213/1307	0.93391	$r=1$	168/1528	1126/1307	0.97146
$r=2$	289/1528	1000/1307	0.95425	$r=2$	289/1528	1000/1307	0.95425
$r=3$	320/1528	924/1307	0.91639	$r=3$	318/1528	926/1307	0.91661
				Abalone, Class 1 vs Class 2, D_F, Greedy			
					Type I	Type II	P_{error}

				$r=1$	168/1528	1126/1307	0.97146
				$r=2$	288/1528	1001/1307	0.95436
				$r=3$	318/1528	926/1307	0.91661
Abalone, Class 1 vs Class 3, D_{KL}				Abalone, Class 1 vs Class 3, D_{KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	551/1528	118/1342	0.44853	$r=1$	626/1528	104/1342	0.48718
$r=2$	511/1528	152/1342	0.44769	$r=2$	578/1528	128/1342	0.47365
$r=3$	502/1528	168/1342	0.45372	$r=3$	509/1528	162/1342	0.45383
Abalone, Class 1 vs Class 3, D_{r-KL}				Abalone, Class 1 vs Class 3, D_{r-KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	626/1528	104/1342	0.48718	$r=1$	210/1528	547/1342	0.54504
$r=2$	579/1528	125/1342	0.47207	$r=2$	368/1528	326/1342	0.48376
$r=3$	546/1528	147/1342	0.46687	$r=3$	473/1528	252/1342	0.49733
Abalone, Class 1 vs Class 3, D_{SKL}				Abalone, Class 1 vs Class 3, D_{SKL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	622/1528	104/1342	0.48456	$r=1$	622/1528	104/1342	0.48456
$r=2$	580/1528	127/1342	0.47422	$r=2$	511/1528	151/1342	0.44694
$r=3$	502/1528	169/1342	0.45447	$r=3$	507/1528	172/1342	0.45997
Abalone, Class 1 vs Class 3, D_H				Abalone, Class 1 vs Class 3, D_H, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	567/1528	109/1342	0.4523	$r=1$	567/1528	109/1342	0.4523
$r=2$	540/1528	135/1342	0.454	$r=2$	493/1528	149/1342	0.43367
$r=3$	519/1528	151/1342	0.45218	$r=3$	513/1528	167/1342	0.46017
				Abalone, Class 1 vs Class 3, D_F, Greedy			
					Type I	Type II	P_{error}
				$r=1$	595/1528	105/1342	0.46764
				$r=2$	503/1528	148/1342	0.43947
				$r=3$	511/1528	168/1342	0.45961
Abalone, Class 2 vs Class 3, D_{KL}				Abalone, Class 2 vs Class 3, D_{KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	132/1307	567/1342	0.5235	$r=1$	458/1307	104/1342	0.42792
$r=2$	325/1307	187/1342	0.38801	$r=2$	476/1307	109/1342	0.44541
$r=3$	383/1307	145/1342	0.40109	$r=3$	382/1307	146/1342	0.40107
Abalone, Class 2 vs Class 3, D_{r-KL}				Abalone, Class 2 vs Class 3, D_{r-KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	458/1307	104/1342	0.42792	$r=1$	132/1307	567/1342	0.5235
$r=2$	476/1307	109/1342	0.44541	$r=2$	235/1307	411/1342	0.48606
$r=3$	383/1307	145/1342	0.40109	$r=3$	477/1307	234/1342	0.53932
Abalone, Class 2 vs Class 3, D_{SKL}				Abalone, Class 2 vs Class 3, D_{SKL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	448/1307	104/1342	0.42027	$r=1$	134/1307	550/1342	0.51236
$r=2$	332/1307	193/1342	0.39783	$r=2$	236/1307	404/1342	0.48161
$r=3$	383/1307	144/1342	0.40034	$r=3$	470/1307	232/1342	0.53248

Abalone, Class 2 vs Class 3, D_H				Abalone, Class 2 vs Class 3, D_H, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	377/1307	128/1342	0.38383	$r=1$	377/1307	128/1342	0.38383
$r=2$	430/1307	121/1342	0.41916	$r=2$	427/1307	120/1342	0.41612
$r=3$	382/1307	149/1342	0.4033	$r=3$	421/1307	128/1342	0.41749
				Abalone, Class 2 vs Class 3, D_F, Greedy			
					Type I	Type II	P_{error}
				$r=1$	423/1307	111/1342	0.40635
				$r=2$	315/1307	186/1342	0.37961
				$r=3$	372/1307	149/1342	0.39565
Banknote, D_{KL}				Banknote, D_{KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	38/762	16/610	0.076098	$r=1$	20/762	0/610	0.026247
$r=2$	23/762	0/610	0.030184	$r=2$	20/762	0/610	0.026247
$r=3$	23/762	0/610	0.030184	$r=3$	23/762	0/610	0.030184
Banknote, D_{r-KL}				Banknote, D_{r-KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	20/762	0/610	0.026247	$r=1$	38/762	16/610	0.076098
$r=2$	20/762	0/610	0.026247	$r=2$	23/762	0/610	0.030184
$r=3$	23/762	0/610	0.030184	$r=3$	23/762	0/610	0.030184
Banknote, D_{SKL}				Banknote, D_{SKL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	20/762	0/610	0.026247	$r=1$	20/762	0/610	0.026247
$r=2$	23/762	0/610	0.030184	$r=2$	23/762	0/610	0.030184
$r=3$	23/762	0/610	0.030184	$r=3$	23/762	0/610	0.030184
Banknote, D_H				Banknote, D_H, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	23/762	1/610	0.031823	$r=1$	23/762	1/610	0.031823
$r=2$	23/762	0/610	0.030184	$r=2$	23/762	0/610	0.030184
$r=3$	23/762	0/610	0.030184	$r=3$	23/762	0/610	0.030184
				Banknote, D_F, Greedy			
					Type I	Type II	P_{error}
				$r=1$	20/762	0/610	0.026247
				$r=2$	23/762	0/610	0.030184
				$r=3$	23/762	0/610	0.030184
Climate, D_{KL}				Climate, D_{KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	2/46	82/494	0.20947	$r=1$	2/46	64/494	0.17303
$r=2$	2/46	59/494	0.16291	$r=2$	2/46	66/494	0.17708
$r=3$	2/46	57/494	0.15886	$r=3$	1/46	52/494	0.127
Climate, D_{r-KL}				Climate, D_{r-KL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r=1$	2/46	64/494	0.17303	$r=1$	2/46	82/494	0.20947

$r = 2$	2/46	51/494	0.14672	$r = 2$	2/46	61/494	0.16696
$r = 3$	1/46	47/494	0.11688	$r = 3$	1/46	37/494	0.096638
Climate, D_{SKL}				Climate, D_{SKL}, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r = 1$	2/46	79/494	0.2034	$r = 1$	2/46	79/494	0.2034
$r = 2$	2/46	54/494	0.15279	$r = 2$	2/46	58/494	0.16089
$r = 3$	1/46	37/494	0.096638	$r = 3$	1/46	37/494	0.096638
Climate, D_H				Climate, D_H, Greedy			
	Type I	Type II	P_{error}		Type I	Type II	P_{error}
$r = 1$	2/46	65/494	0.17506	$r = 1$	2/46	65/494	0.17506
$r = 2$	1/46	51/494	0.12498	$r = 2$	2/46	57/494	0.15886
$r = 3$	1/46	48/494	0.11891	$r = 3$	2/46	35/494	0.11433
				Climate, D_F, Greedy			
					Type I	Type II	P_{error}
				$r = 1$	2/46	72/494	0.18923
				$r = 2$	1/46	50/494	0.12295
				$r = 3$	2/46	35/494	0.11433