

TOMomentum因子报告

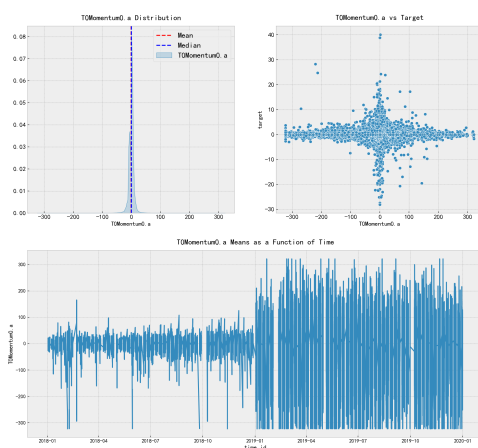
1. 描述性统计

数据统计性描述，分别计算了特征分布的均值、标准差、分位数、偏度、峰度以描述特征的分布情况：

index	count	mean	std	min	25%	50%	75%
TOMomentum0. a	2302740.0	-0.349	12.908	-3221.501	-4.575	0.0	4.349
TOMomentum0. b	2306678.0	0.0	0.008	-0.625	-0.002	-0.0	0.002
TOMomentum0. c	2306677.0	-0.0	0.001	-0.11	-0.0	-0.0	0.0
TOMomentum1. a	2302044.0	-0.805	16.551	-872.937	-6.94	0.0	5.992
TOMomentum1. b	2306573.0	0.0	0.002	-0.12	-0.001	-0.0	0.001
TOMomentum1. c	2306572.0	0.0	0.0	-0.001	-0.0	-0.0	0.0
TOMomentum2. a	2300876.0	-1.23	19.699	-949.711	-8.959	0.0	8.03
TOMomentum2. b	2306443.0	0.0	0.001	-0.061	-0.001	-0.0	0.0
TOMomentum2. c	2306442.0	0.0	0.0	-0.0	-0.0	-0.0	0.0

index	max	n_missing	n_infinite	median	skew	kurtosis
TOMomentum0. a	372.636	3962.0	0.0	0.0	-12.727	1913.949
TOMomentum0. b	0.597	24.0	0.0	-0.0	0.574	609.685
TOMomentum0. c	0.161	25.0	0.0	-0.0	14.145	4632.716
TOMomentum1. a	373.286	4658.0	0.0	0.0	-4.13	112.836
TOMomentum1. b	0.054	129.0	0.0	-0.0	-2.176	171.079
TOMomentum1. c	0.001	130.0	0.0	-0.0	10.931	3393.72
TOMomentum2. a	373.016	5826.0	0.0	0.0	-3.51	75.429
TOMomentum2. b	0.029	259.0	0.0	-0.0	-1.827	106.486
TOMomentum2. c	0.0	260.0	0.0	-0.0	8.703	2571.22

通过图形可视化描述数据的数值分布，
与预测目标收益率TickReturn的相关性， 以及特征在时序上的表现：



2. 相关系数

分别计算给定不同参数和不同计算方式的特征与多个预测目标的相关系数，直接观察其线性关系：

index	TickReturn_Rank	TickReturnLabelThres_Rank
TOMomentum0. a	141	152
TOMomentum0. b	151	145
TOMomentum0. c	118	123
TOMomentum1. a	171	168
TOMomentum1. b	150	157
TOMomentum1. c	145	147
TOMomentum2. a	170	169
TOMomentum2. b	120	120
TOMomentum2. c	156	155

3. t检验

纳入全部的特征，分别对不同预测目标进行线性回归，观察其系数的显著性表现，可以一定程度上判断线性关系，同时也可以帮助判断同一大类特征内不同参数的取舍：

index	TickReturn_Rank	TickReturnLabelThres_Rank
TOMomentum0. a	93	75
TOMomentum0. b	45	35

TOmomentum0. c	180	182
TOmomentum1. a	28	27
TOmomentum1. b	20	22
TOmomentum1. c	169	172
TOmomentum2. a	39	39
TOmomentum2. b	11	10
TOmomentum2. c	46	38

4. F检验

纳入全部的特征，分别对不同预测目标进行线性回归，对某一大类的特征进行F检验，从而判断该类特征是否有线性预测能力：

index	TickReturn_Rank	TickReturnLabelThres_Rank
TOmomentum. a	24	22
TOmomentum. b	12	11
TOmomentum. c	14	13

5. 熵-信息增益

分别将连续的特征进行离散化，分为一百个标签，判断给定某一特征后，能给模型预测目标的熵带来的信息增益情况，由于TickReturn为连续型变量，因而通过直接按正负性划分为两类判断：

index	TickReturn_Rank	TickReturnLabelThres_Rank
TOmomentum0. a	82	96
TOmomentum0. b	65	71
TOmomentum0. c	102	115
TOmomentum1. a	60	66
TOmomentum1. b	67	58
TOmomentum1. c	81	87
TOmomentum2. a	64	65
TOmomentum2. b	73	67
TOmomentum2. c	76	80

6. 分组指标评价

将某大类因子放入模型，分别根据评价指标（R_square, F1_score)获取该类因子排名：

index	TickReturn_Rank	TickReturnLabelThres_Rank
TO Momentum. a	16	25
TO Momentum. b	24	19
TO Momentum. c	18	31

7. 模型SHAP

通过SHAP衡量各个特征的表现：

index	TickReturn_Rank	TickReturnLabelThres_Rank
TO Momentum0. a	75	73
TO Momentum0. b	5	22
TO Momentum0. c	82	78
TO Momentum1. a	147	177
TO Momentum1. b	6	19
TO Momentum1. c	183	184
TO Momentum2. a	133	182
TO Momentum2. b	21	62
TO Momentum2. c	163	185

10. 评分

特征的评分分为以下两部分： 线性部分和非线性部分。线性部分考纳线性回归的统计性结果，对相应的排名进行映射为分数，最后将所有线性评分进行平均处理得到线性分数。非线性部分则参考了熵，SHAP，LGBM重要性等指标排名，同样进行映射平均。

10.1 线性评分

index	TickReturn	TickReturnLabelThres
TO Momentum0. a	68.54838709677419	73.9247311827957

TO Momentum0. b	84. 67741935483872	87. 63440860215053
TO Momentum0. c	47. 84946236559139	47. 58064516129032
TO Momentum1. a	86. 0215053763441	86. 82795698924731
TO Momentum1. b	91. 3978494623656	91. 12903225806451
TO Momentum1. c	50. 806451612903224	50. 26881720430108
TO Momentum2. a	83. 06451612903226	83. 6021505376344
TO Momentum2. b	93. 81720430107528	94. 35483870967744
TO Momentum2. c	83. 87096774193549	86. 29032258064515

10.2 非线性评分

index	TickReturn	TickReturnLabelThres
TO Momentum0. a	68. 9964157706093	65. 23297491039426
TO Momentum0. b	83. 15412186379929	79. 92831541218636
TO Momentum0. c	63. 79928315412186	59. 856630824372765
TO Momentum1. a	60. 03584229390682	51. 97132616487455
TO Momentum1. b	82. 61648745519713	82. 79569892473118
TO Momentum1. c	49. 46236559139785	45. 87813620071685
TO Momentum2. a	61. 82795698924732	51. 25448028673835
TO Momentum2. b	78. 85304659498209	73. 47670250896059
TO Momentum2. c	53. 9426523297491	46. 95340501792114

10.3 综合评分

index	TickReturn	TickReturnLabelThres
TO Momentum0. a	68. 772	69. 579
TO Momentum0. b	83. 916	83. 781
TO Momentum0. c	55. 824	53. 719
TO Momentum1. a	73. 029	69. 4
TO Momentum1. b	87. 007	86. 962
TO Momentum1. c	50. 134	48. 073
TO Momentum2. a	72. 446	67. 428
TO Momentum2. b	86. 335	83. 916
TO Momentum2. c	68. 907	66. 622