Some results on spatial data

Yi Li

November 3, 2016

— CyberSEES

---

1. PCA for image reconstruction

1.1 PCA reconstruction formula

In PCA, we have:

- the data matrix $\mathbb{X}_{n\times p}$, where n is the sample size, and p is the dimension of the data.
- sample covariance $S = \frac{1}{n-1}\mathbb{X}^\top\mathbb{X}$ [1].
- eigen decomposition of $S = V_{p\times p}\Lambda_{p\times p}V_{p\times p}^\top$.
  - columns of V: eigenvectors of V, principal axes(direction), loadings
- projection of data on principle axes – principle components (PC scores): $\mathbb{Z} = \mathbb{X}_{n\times p}V_{p\times p}$
  - i^{th} row of $\mathbb{Z}$: coordinates of the i^{th} data point in the new PC space
  - j^{th} column of $\mathbb{Z}$: j^{th} PC
  - reconstruct $\hat{\mathbb{X}} = \mathbb{Z}V^\top = \mathbb{X}VV^\top$
- if choose first d columns of V, we get $V_{p\times d}$:
  - $\mathbb{Z}_{n\times d} = \mathbb{X}_{n\times p}V_{p\times d}$
  - reconstruct $\hat{\mathbb{X}}_{n\times p} = \mathbb{Z}_{n\times d}V_{d\times p}^\top = \mathbb{X}_{n\times p}V_{p\times d}V_{d\times p}^\top$
  - $VV^\top$ is projection matrix
- above assumes the mean is substracted from the data matrix
  - $\mathbb{Z} = (\mathbb{X} - \mu)V$, $\mathbb{X} = \mathbb{Z}V^\top + \mu$
- in the validation setting, $\hat{\mathbb{X}}_{n_{test}\times p} = \mathbb{X}_{n_{test}\times p}V_{p\times d}V_{d\times p}^\top$
  - V is calculated from the training data
  - $\mathbb{X}_{n_{test}\times p}$ is from testing data

PCA reconstruction = PC score × eigenvectors$^\top$+mean

[1] true covariance Σ

1.2 Simulated data[2]

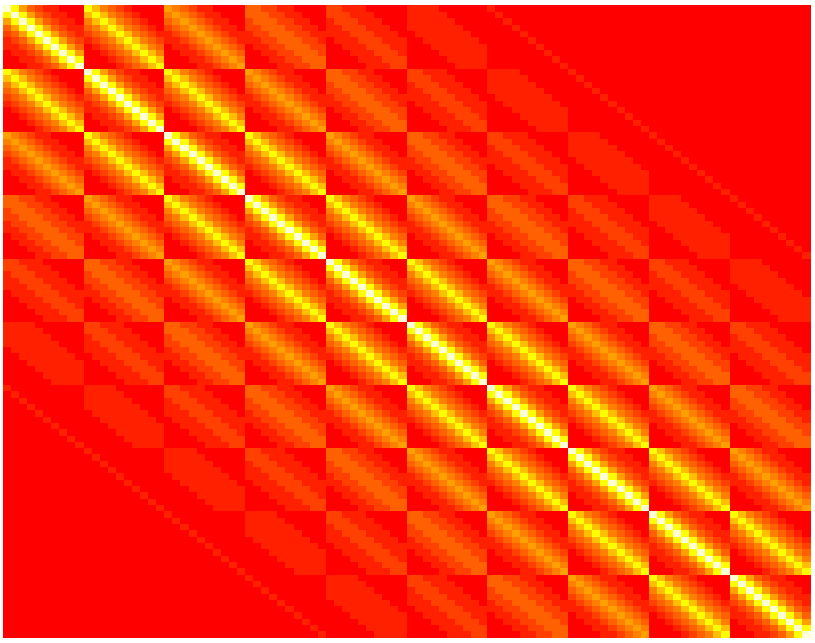
- Assume the spatial data are correlated acording to their spatial distance.
- Then we have the covariance structure as "block bandable"
- Modify the sample covariance matrix: keep the central diagonal elements in each block, set small off-diagonal elements to zero.
- Here are some simulation in R [R Core Team, 2015]

[2] with the block-bandable matrices

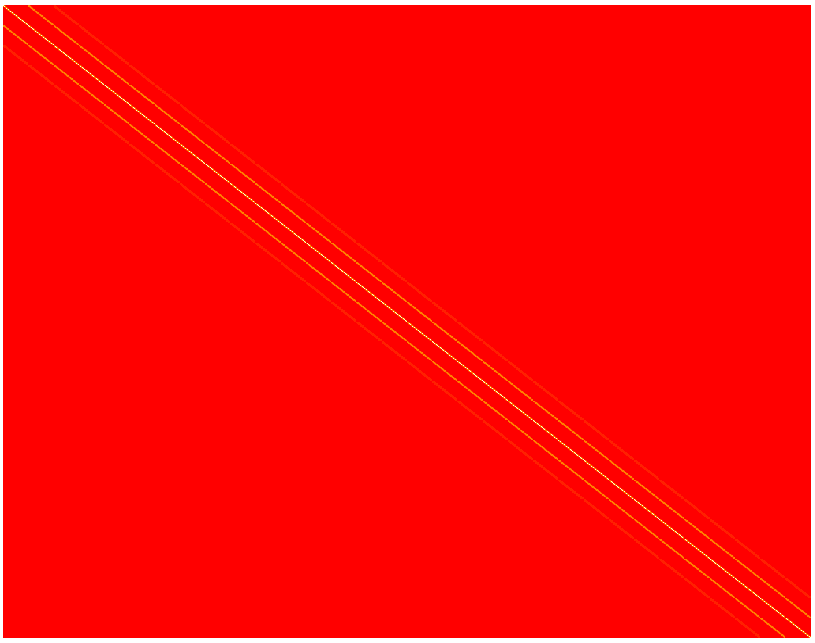
X: 2D Spatial Stations Labels

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

S: Covariance Matrix Structure



S: Covariance Matrix Structure

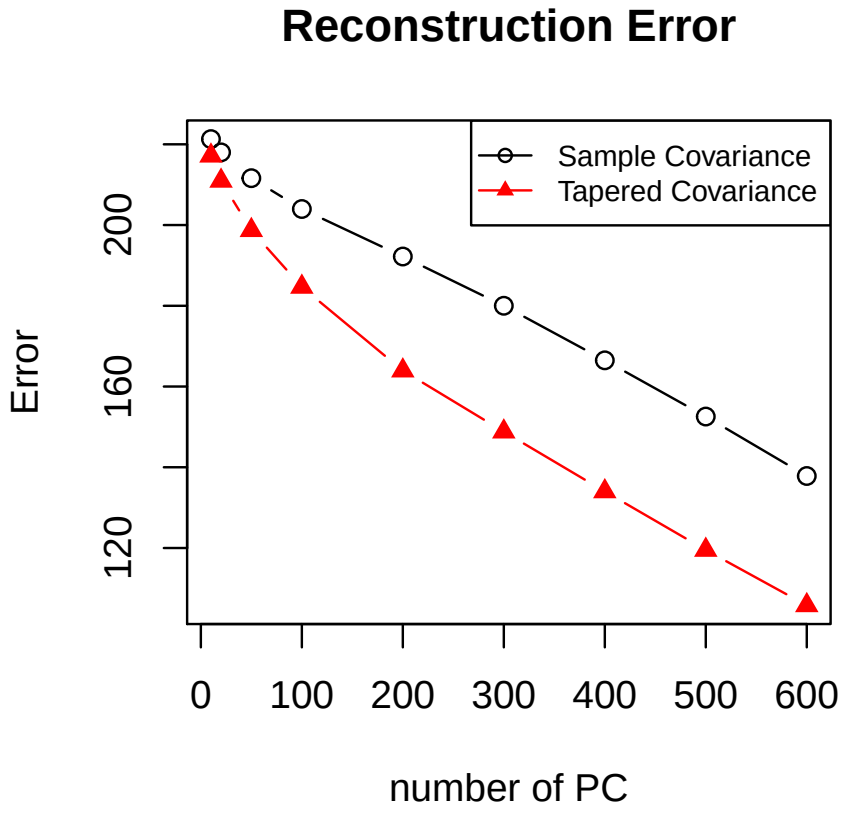


```
dim(X0) # training
```

```
## [1]   80 1024
```

```
dim(X1) # testing
```

```
## [1]   20 1024
```



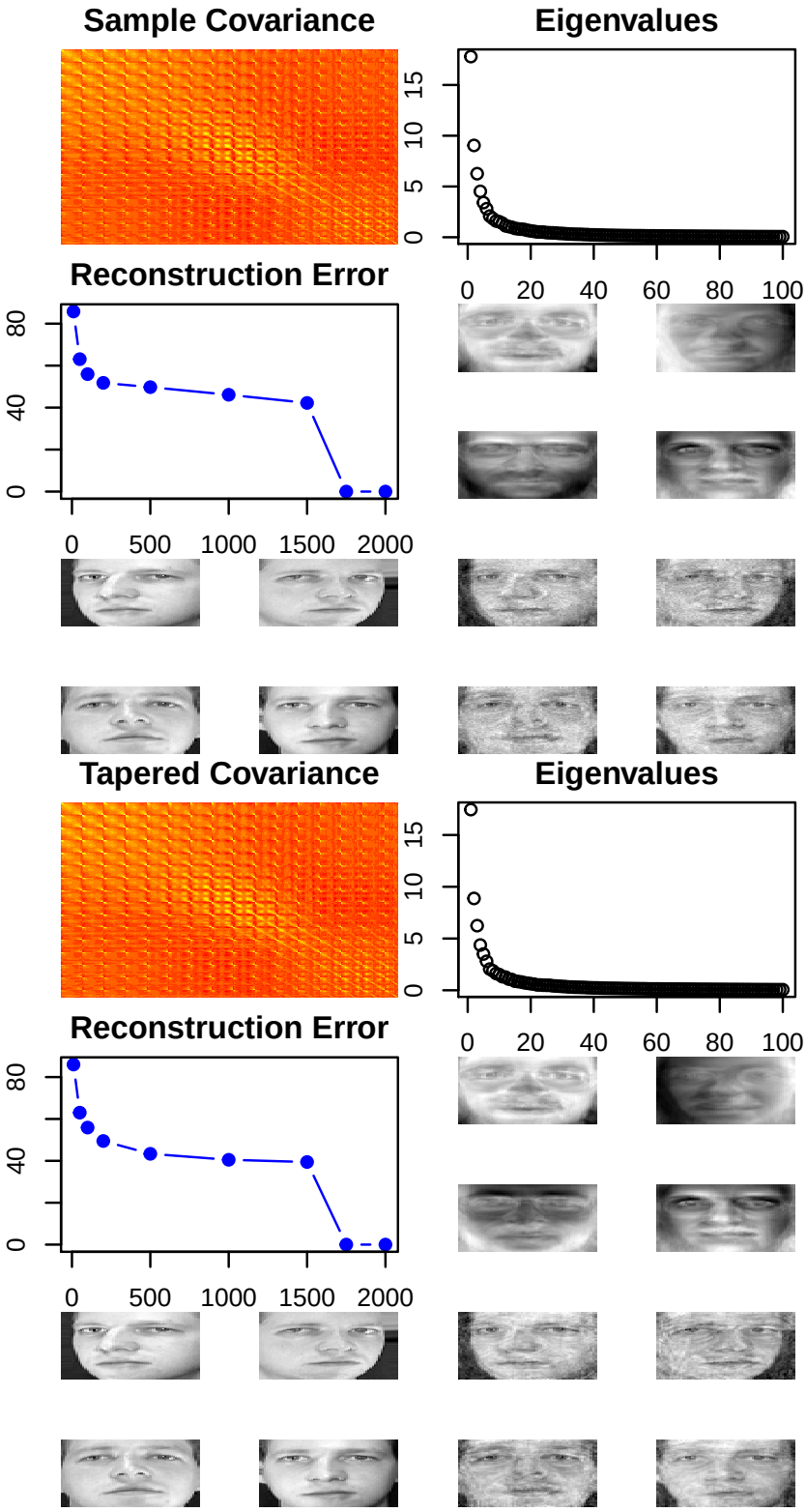
1.3 Face data

```
dim(X0) # training
```

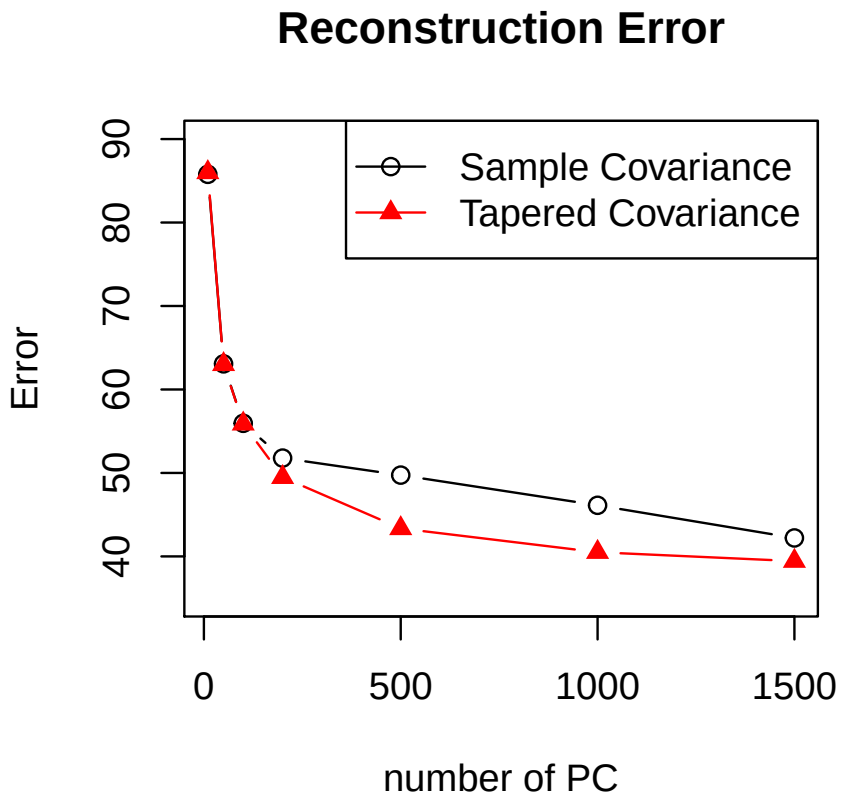
```
## [1]  160 4096
```

```
dim(X1) # testing
```

```
## [1]  240 4096
```



- Sample Covariance Structure
- Eigenvalues (first 100)
- Reconst. Error on test
- Eigenfaces * 4
- Test image
- Reconstructed Test image (1k PCs)



References

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.R-project.org/`.