Homework Exercises

Mladen Victor WICKERHAUSER

PMF 2022: Dimensionality Reduction and Manifold Estimation

July 26, 2022

If an exercise consists of just a statement, then prove the statement.

1 Manifolds

Exercise 1: $|\langle x, y\rangle| \leq \|x\|\|y\|$. When is there equality?

Exercise 2:

(a) Second countable implies first countable.

(b) Second countable implies separable.

Exercise 3:

(a) A metric space is a first countable Hausdorff topological space.

(b) A separable metric space is second countable.

Exercise 4: (Lindelöf) If X is a separable metric space, and $A \subset X$ is any subset, then every open cover of A has a countable subcover.

Exercise 5: Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. If $\phi : X \to Y$ is bijective and continuous, then for each $G_Y \in \mathcal{T}_Y$ there exists $G_X \in \mathcal{T}_X$ such that $\phi(G_X) = G_Y$.

Exercise 6: (G, ϕ) is differentially compatible with (G, I) iff $\phi : \mathbf{E}^d \to \mathbf{E}^d$ is differentiable on G. (Hint: you may use the Inverse Function Theorem.)

Exercise 7: Newton-Raphson K' iteration converges to the same unique root x' as contraction iteration K.

Exercise 8: Linear $\mathbf{F}$ implies $D_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \mathbf{y})$ and $D_{\mathbf{y}}\mathbf{F}(\mathbf{x}, \mathbf{y})$ are constant.

Exercise 9: Show that ξ is well-defined. Namely, explain why the directional derivatives of f agree for all of $v's$ equivalent curves through x.

Exercise 10: $G(x)$, the set of all germs at x, is an algebra under pointwise addition and multiplication.

2 Regression

Exercise 11: Unique solution $\{a_j\}$ exists for each $\{y_i : 0 \leq i \leq n\}$ iff matrix $\{f_j(x_i) : 0 \leq i, j \leq n\}$ is nonsingular.

Exercise 12: Distinct x_i's make Vandermonde matrix V invertible.

Exercise 13: If $\hat{\mathbf{p}}$ is linear and $\mathbf{y}$ has finite expectation and variance, then $\hat{\mathbf{p}}(\mathbf{y})$ has finite variance and expectation.

Exercise 14: K, K' convex implies $K \cap K'$ is convex.

Exercise 15: If $\mathbf{x}$ is in the 2-simplex defined by $\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_c$, then the convex combination $\mathbf{x} = p_a\mathbf{x}_a + p_b\mathbf{x}_b + p_c\mathbf{x}_c$ is unique.

Exercise 16: Complete the proof of the Delaunay tesselation theorem. Hint: see K.Q.Brown (1979).

Exercise 17: If S is in general position, then no $x \in \mathcal{M}$ can be equidistant from 4 or more centers.

Exercise 18: The dual of the dual of (V, E, F) is isomorphic to (V, E, F).

3 Compression

Exercise 19: $\operatorname{tr}\bar{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n} \|\mathbf{x}_i - \bar{\mathbf{x}}\|^2$, namely, the trace of the covariance matrix is the total sample variance.

Exercise 20: Transform coding gain $G(U)$ is maximized by the empirical K-L basis U for $\mathbf{x}$. (Hint: use Cholesky factorization.)

Exercise 21: The Haar scaling expansion P_0 is an orthogonal projection.

Exercise 22: $HH^*x = x$, $GG^*x = x$, and $x = H^*Hx + G^*Gx$.

4 Reduction

There are no exercises for this section. Download and install the R programming system and experiment with the files

```
04clust.txt
04datav.txt
04isomap.txt
04malda.txt
04stepr.txt
04trees.txt
```

5 Perron-Frobenius Theorem

Exercise 23: M is irreducible, namely M is nonnegative and $\exp(M) - I$ is positive, if and only if $(\forall i, j)(\exists k) M^k(i, j) > 0$.

Exercise 24: Prove Gershgorin's theorem:

Theorem. *Suppose M is an $n \times n$ matrix over $\mathbf{C}$. For $i = 1, \ldots, n$, define the Gershgorin disc $G_i \subset \mathbf{C}$ by*

$$G_i \stackrel{\text{def}}{=} \Big\{ z \in \mathbf{C} : |z - M(i,i)| \le \sum_{j \ne i} |M(i,j)| \Big\}.$$

Then every eigenvalue of M lies in $\bigcup_i G_i$.

6 Diffusion Maps

Exercise 25: If P is row stochastic, then $\|P^k\|_\infty = 1$ for all k.

Exercise 26: Suppose that $P = USV^T$ is a singular value decomposition of the $n \times n$ matrix P, where U, V are orthogonal, and S is diagonal. Then

$$P = \sum_l s_l \mathbf{u}_l \otimes \mathbf{v}_l \quad \text{meaning} \quad P(i,j) = \sum_l s_l \mathbf{u}_l(i) \mathbf{v}_l(j).$$

Exercise 27: Let P be the transition matrix obtained from the gaussian kernel matrix on a n-point data set. Let Ψ and Φ be the biorthogonal duals derived from P

(a) $P = \Psi \Lambda \Phi^T$, with $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n. \end{pmatrix}$.

(b) Column ψ_l of Ψ is an eigenvector of P with eigenvalue λ_l.

(c) $P^k = \Psi \Lambda^k \Phi^T$.

7 Markov Chain Monte Carlo

Exercise 28: Let $\mathbf{p} = (p_1, \ldots, p_k)$ be a Dirichlet posterior for a Dirichlet prior and multinomial experimental likelihood, so

$$\left(p_1^{\alpha_1 - 1} \cdots p_k^{\alpha_k - 1}\right) \left(p_1^{n_1} \cdots p_k^{n_k}\right) \propto p_1^{n_1 + \alpha_1 - 1} \cdots p_k^{n_k \alpha_k - 1},$$

Fix i with $1 \le i \le k$. Under what conditions $\mathrm{Var}(p_i) \to 0$ as the number of experiments tends to infinity?

Exercise 29: Suppose that $\{X_k : k = 1, 2, \ldots\}$ is an ergodic Markov chain on a finite state space $S = \{1, \ldots, m\}$. Let $\pi \in \mathbf{R}^m$ be its stationary distribution. Then for any bounded function F on S,

$$\frac{1}{N} \sum_{k=1}^{N} F(X_k) \to \sum_{i=1}^{m} F(i) \pi(i),$$

almost surely as $N \to \infty$.