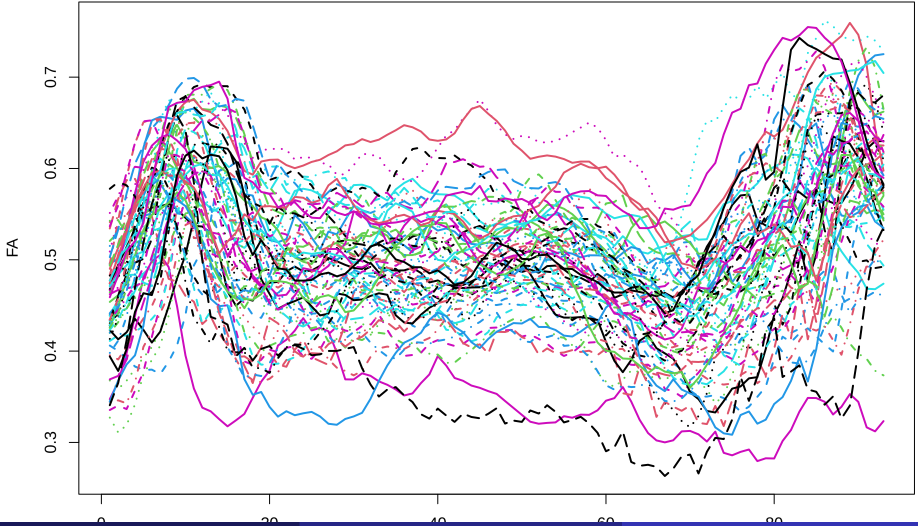
# Functional Principal Components Analysis

School of Statistics and Data Science

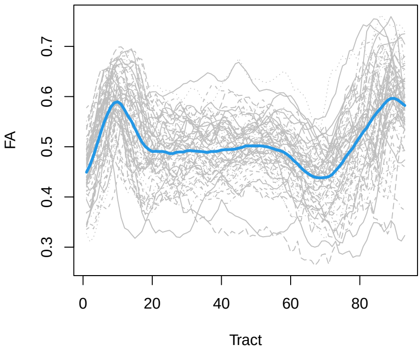
# Diffusion Tensor Imaging Example

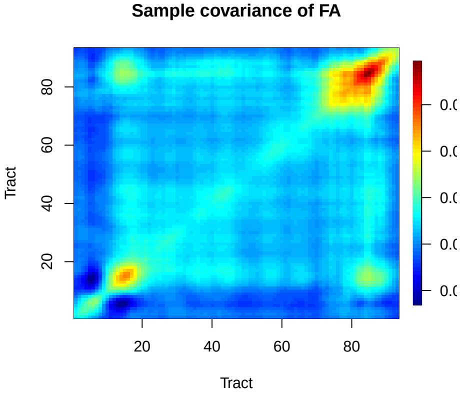
Fractional anisotropy (FA) is a measure of water diffusion in the brain. We consider FA tract profiles for the corpus callosum (CCA) in MS patients.



# Sample Mean and Covariance

$$\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n} X_i(t), \quad \hat{c}(s,t) = \frac{1}{n-1}\sum_{i=1}^{n}\{X_i(s) - \hat{\mu}(s)\}\{X_i(t) - \hat{\mu}(t)\}$$



# Review of PCA

Principal Component Analysis (PCA) [Karl Pearson, 1901] is statistical technique aimed at linear dimension reduction in multivariate analysis

- Capture the main modes of variation
- Reduce the dimensionality
- Extract low-dimensional features

Functional principal component analysis (fPCA) [Rao 1958] extends the ideas to the case when data are curves"

# Review of PCA

- Directions of greatest variation
- Dimension reduction-subspace closest to the data
- Frequently picks out interpretatble contrasts

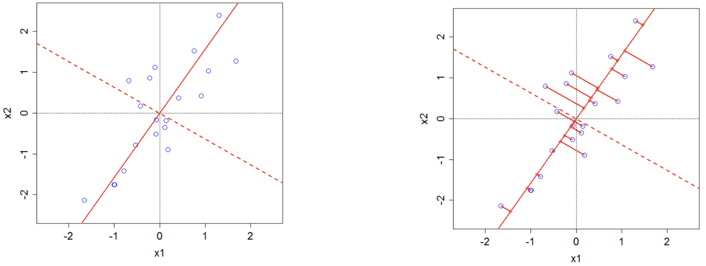



Figure 1: PCA illustration

# A little analysis

## Total Variation

Measure total variation in the data as total squared distance from center:

$$\sum_{j=1}^{p}\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2 = \text{trace}(\Sigma)$$

## Variance of Projection

If $X$ has covariance $\Sigma$, the variance of $\gamma^\top X$ is $\gamma^\top \Sigma \gamma$.

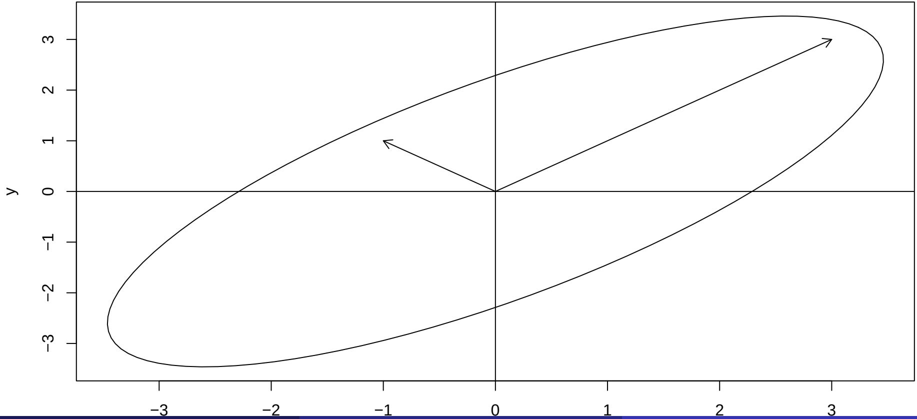
## Optimization Problem

To maximize $\gamma^\top \Sigma \gamma / \gamma^\top \gamma$, we solve the eigen-equation:

$$\Sigma\gamma = \lambda\gamma.$$

# Review of PCA

- Let $\lambda_1 > \cdots > \lambda_p$ be eigenvalues of $\Sigma = var(X)$, with $\gamma_1, \ldots, \gamma_p$, corresponding eigenvectors
- The direction cosine of the $i$-th principal axis is $\gamma_i$
- The length of the $i$-th principal semi axis is $c\lambda_i$



# Mechanics of PCA

- Estimate covariance matrix: $\hat{\Sigma} = \frac{1}{n}\sum_i (X_i - \bar{X})(X_i - \bar{X})^\top$
- Take the eigen-decomposition of $\Sigma = \Gamma\Lambda\Gamma^\top$
- Columns of $\Gamma$ are orthogonal, represent a new basis
- $\Lambda$ is diagonal, entries give variances of data along corresponding directions $\Gamma$

$$\text{Proportion of variance explained: } \lambda_k / \sum_k \lambda_k$$

- Order $\Lambda, \Gamma$ in terms of decreasing $\lambda_i$
- $\gamma_i$ is the $i$-th column of $\Gamma$ contains the principal component lodgings
- From original data $X$, $(X - \bar{X})^\top \gamma_i$ is the $i$-th principal component score, co-ordinate in new basis.

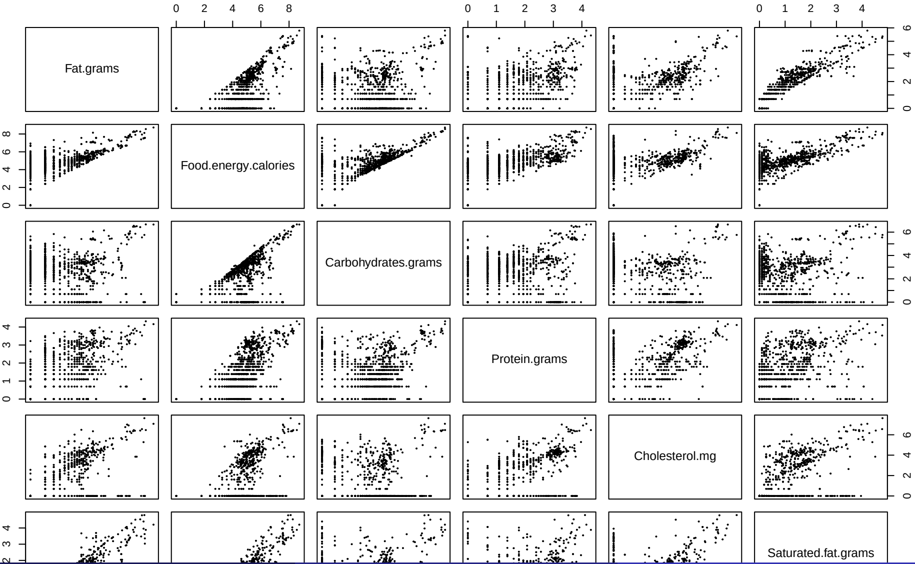
# Properties of Principal Components

The $p$-principal components of $X$, obtained through the transformation $Y = \Gamma'(X - \mu)$ have the following properties

- $E(Y) = 0$
- $Cov(Y) = \Lambda$
- $Var(Y_1) \geq Var(Y_2) \geq \cdots \geq Var(Y_p) \geq 0$
- $\sum_i Var(Y_i) = tr\,\Sigma$
- $\prod_i Var(Y_i) = |\Sigma|$

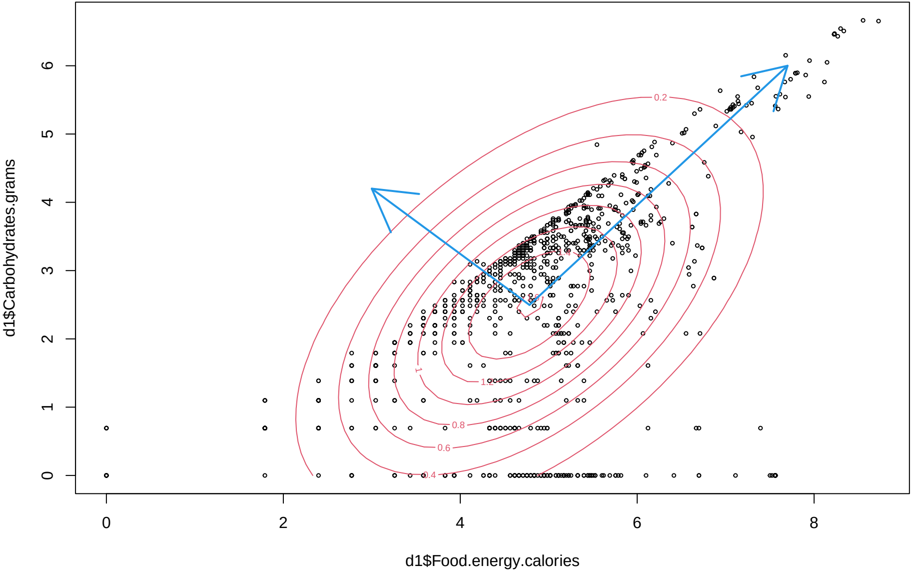
# Example: PCA in Nutrition Value

Nutritional data from 961 food items are listed alphabetically



# Ellipsoids and Information

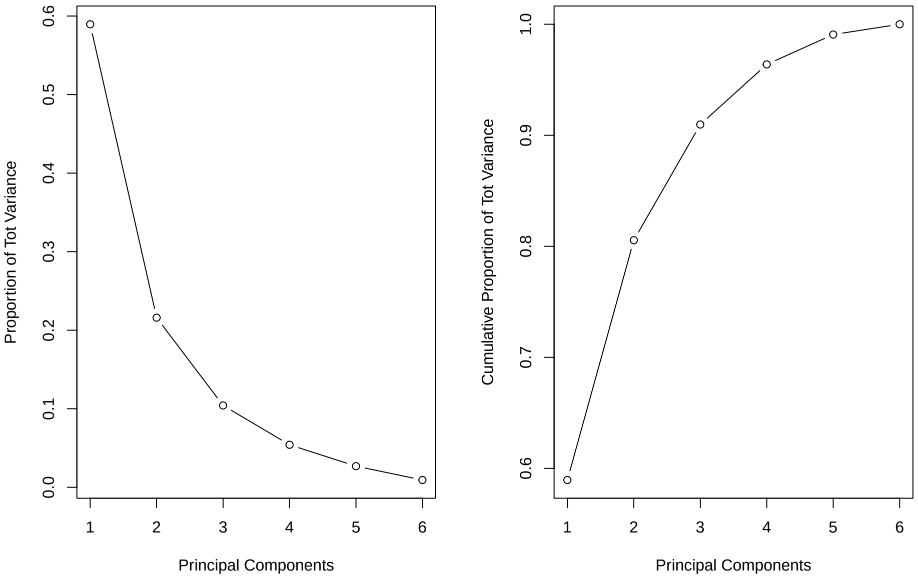
Consider the relationship between carbohydrates and calories



**PC weights $\Gamma$**

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 |
|---|---|---|---|---|---|---|
| Fat.grams | 0.48 | 0.17 | -0.39 | 0.18 | 0.27 | 0.70 |
| Food.energy.calories | 0.44 | -0.38 | -0.17 | 0.13 | -0.78 | -0.04 |
| Carbohydrates.grams | 0.14 | -0.81 | 0.03 | -0.40 | 0.39 | 0.07 |
| Protein.grams | 0.41 | -0.08 | 0.68 | 0.55 | 0.22 | -0.11 |
| Cholesterol.mg | 0.39 | 0.36 | 0.43 | -0.70 | -0.16 | 0.12 |
| Saturated.fat.grams | 0.48 | 0.18 | -0.41 | -0.07 | 0.29 | -0.69 |

# PCA Analysis of Nutrition Data



# Projections in $L^2$

- For $k \geq 1$, let $\{u_1, \ldots, u_k\}$ be an orthonormal basis in $L^2$

$$\begin{array}{lll} \langle u_j, u_{j'} \rangle = 0 & \text{if} & j \neq j' \\ \langle u_j, u_j \rangle = 1 & \text{for all} & j = 1, \ldots, k \end{array}$$

- The projection of a random function $X$ onto the space spanned by $\{u_1, \ldots, u_k\}$ is

$$\sum_{j=1}^{k} \langle X, u_j \rangle \, u_j = \sum_{j=1}^{k} \left\{ \int X(t) \, u_j(t) \; dt \right\} u_j$$

- The expected distance between $X$ and its projection in the span of $\{u_1, \ldots, u_k\}$ is

$$S_X(u_1, \ldots, u_k) = E \left|\left| X - \sum_{j=1}^{k} \langle X, u_j \rangle \, u_j \right|\right|^2 = E||X||^2 - \sum_{j=1}^{k} \langle X, u_j \rangle^2$$

# Mercer and Karhunen-Loeve Revisited

**Mercer's Theorem**: Let $X(t)$ be a square integrable random function with covariance function $c(s,t)$. There exists an orthonormal basis $\{\phi_j\}_{j\geq 1}$ of continuous eigenfunctions in $L^2(T)$, and eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots > 0$ s.t.

$$c(s,t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s)\phi_j(t)$$

with $\sum_j^{\infty} \lambda_j = \int c(t,t)\,dt < \infty$

**Karhunen-Loeve Theorem**: Let $X(t)$ be a square integrable function, with mean $\mu(t)$ and covariance function $c(s,t)$. We have

$$X(t) = \mu(t) + \sum_{j=1}^{\infty} \nu_j \phi_j(t)$$

with $E(\nu_j) = 0$, $E(\nu_j^2) = \lambda_j$, and $E(\nu_j, \nu_{j'}) = 0$.

# KL Representation and Optimal Projections

**KL Thm. restated**: Let $X$ be a square integrable random function with mean $\mu$, and covariance $c$ with eigenfunctions $\{\phi_j\}_{j\geq 1}$ and eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots > 0$.

The expected projection distance $S_X(u_1, \ldots, u_k)$ is minimized by setting $u_j = \phi_j$, $j = 1, \ldots, k$.

For $k \to \infty$ it follows that

$$X - \mu = \sum_{j=1}^{\infty} \langle X - \mu, \phi_j \rangle \, \phi_j = \sum_{j=1}^{\infty} \nu_j \, \phi_j$$

- $\nu_j = \langle X - \mu, \phi_j \rangle$ is a random variable, wereas $\phi_j$ is interpreted as an unknown fixed function.

# Functional PCA

- Instead of covariance matrix $\Sigma$, we have a bivariate function $c(s,t)$.
- Re-interpret eigen-decomposition

$$\Sigma = \Gamma \Lambda \Gamma^\top = \sum_j \lambda_j \gamma_j \gamma_j^\top$$

- For functions, this is the decomposition

$$c(s,t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \phi_j(t)$$

- The $\lambda_j$ represents amount of variation in direction $\phi_j(t)$

# Re-interpretation

- Collection of $\{X_i\}_{i=1}^n$
- Find the direction $\phi_1(t)$ that maximizes

$$Var\left[\int_{\mathcal{T}} \phi_1(t) X_i(t) dt\right], \quad s.t. \quad \int_{\mathcal{T}} \phi_1^2(t) dt = 1.$$

- For the second direction, maximize the variance subject to the orthogonality condition

$$Var\left[\int_{\mathcal{T}} \phi_2(t) X_i(t) dt\right], \quad s.t. \quad \int_{\mathcal{T}} \phi_2^2(t) dt = 1, \int_{\mathcal{T}} \phi_1(t)\phi_2(t) dt = 0$$

# PCA and KL

The bivariate covariance can be decomposed

$$c(s,t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s)\phi_j(t)$$

with $\{\phi_j\}$ orthogonal.

- The $\{\phi_j\}$ are the principal components, successively maximize $Var[\int_{\mathcal{T}} \phi_j(t) X_i(t) dt]$
- $\lambda_j$ gives this variance
- $\lambda_j / \sum_j \lambda_j$ is the proportion of variance explained
- $\{\phi_j\}$ is a basis system, $X_i(t) = \sum_{j=1}^{\infty} \nu_{ij}\phi_j(t)$
- Principal component scores are

$$\nu_{ij} = \int_{\mathcal{T}} [X_i(t) - \bar{X}(t)]\phi_j(t)dt$$

# KL and Variance Decomposition

- Kl represents $X(t)$ through a one-to-one map $X(t) \to (\nu_1, \nu_2, \ldots)^T$
- $\lambda_j$ is the average variance of the j-th fPC $\phi_j$,

$$\int \mathrm{Var}\{\nu_j \phi_j(t)\} dt = \mathrm{Var}(\nu_j) \int \phi_j(t)^2 dt = \lambda_j$$

- $E||X - \mu||^2 = \sum_{j=1}^{\infty} \lambda_j$
- Percentage explained variance (PEV) by the first $k$ fPCs

$$\frac{\sum_{j=1}^{k} \lambda_j}{\sum_{j=1}^{\infty} \lambda_j}$$

# Estimation of fPCs

The literature on fPCs estimation from data distinguishes the following cases

1. Estimation of fPCs from true iid curves
2. Estimation of fPCs from densely sampled curves
3. Estimation of fPCs from sparsely sampled curves

Case (1) is often of pure scholastic interest, even though not impossible in rare applications

Case (2) and (3) are often reliant on a subjective assessment of what is meant by "sparse''

# Estimation fPCs (true iid curves)

Observe $X_1, \ldots, X_n$ iid in $L^2(T)$

1. Estimate $\hat{\mu}$ and $\hat{c}$ (sample mean and covariance)
2. For any finite truncation $k < \infty$, solve the eigenequations

$$\int \hat{c}(s,t)\hat{\phi}_j(s)ds = \hat{\lambda}_j \hat{\phi}_j(t), \quad \text{for } j = 1, \ldots, k$$

   Typically this is practically obtained through the eigenanalysis of $\hat{c}(\mathbf{t}) \in \mathbb{R}^{m \times m}$ evaluated over a dense discrete grid of values $\mathbf{t} = (t_1, \ldots, t_m)^T$

3. Estimate fPC scores

$$\hat{\nu}_{ij} = \int \{X_i(t) - \hat{\mu}(t)\}\hat{\phi}_j(t)\, dt$$

   This is often achieved through numerical quadrature

# Sample Variance Decomposition

Observe $X_1, \ldots, X_n$ iid in $L^2(T)$.

- For any $g \in L^2$, the statistic

$$\frac{1}{n}\sum_{i=1}^{n} \langle X_i - \hat{\mu}, g \rangle^2 = \left\langle \hat{C}(g),\, g \right\rangle$$

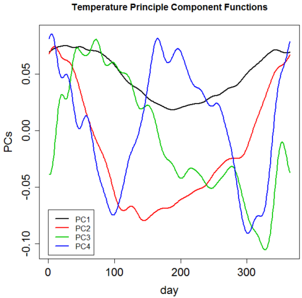
  can be viewed as the sample variance along a function $g$, or in the direction of $g$.

- Considering $g = \hat{\phi}_j$, we have

$$\sum_{j=1}^{n} \frac{1}{n}\sum_{i=1}^{n} \left\langle X_i - \hat{\mu},\, \hat{\phi}_j \right\rangle^2 = \frac{1}{n}\sum_{i=1}^{n} ||X_i - \hat{\mu}||^2 = \sum_{j=1}^{n} \hat{\lambda}_j$$

  Thus we say that $\lambda_j$ explains a fraction of the total sample variance in the direction of $\hat{\phi}_j$

# Example: Canadian Temperature Data

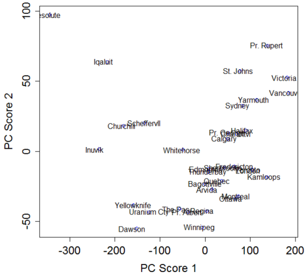



- PC1 – over-all temperature
- PC2 – relative temperature of winter and summer
- PC3 – contrast between fall and spring
- PC4 – relative lengths of summer/winter

Figure 2: PCs of Canadian Temperature Data

# Example: Canadian Temperature Data

Sanity check: we can plot the first two PC scores for each observation.




- First PC: over-all temperature.
- Second PC: contrast between Summer and Winter.

Figure 3: Scores of Canadian Temperature Data

# Theoretical Properties of Sample fPCs

Assume $X_1, \ldots, X_n$ are iid with the same distribution as $X$ with mean $\mu$ and covariance function $c$

- If $E||X||^2 < \infty$, then the sample mean $\hat{\mu}$ is unbiased $E(\hat{\mu}) = \mu$ and mean square consistent $E||\hat{\mu} - \mu|| = O(n^{-1})$
- If $E||X||^4 < \infty$, the the sample covariance $\hat{c}$ is unbiased and mean square consistent
- If $\lambda_1 > \lambda_2 > \cdots \geq 0$, and $\hat{\varsigma}_j = \text{sign} \int \hat{\phi}_j(t)\phi_i(t)dt$, then

$$\limsup_{n\to\infty} n\, E||\hat{\varsigma}\hat{\phi}_j - \phi_j||^2 < \infty, \quad \text{and} \quad \lim_{n\to\infty} n\, E\{|\hat{\lambda}_j - \lambda_j|^2\} < \infty$$

# Estimation of FPCs (dense sampling)

- Observed data $\{Y_i(t_{ij}),\ j = 1, \ldots, m_i\}_{i=1}^{n}$ iid
- Assume

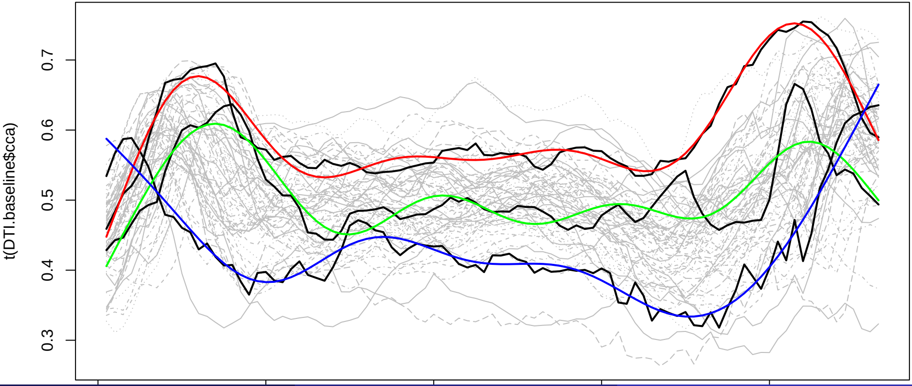
$$Y_i(t_{ij}) = X_i(t_{ij}) + \epsilon_{ij};\ \ X_i \in L^2[T],\ \ \epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

- Common approach:

1. Obtain $\hat{X}_i$ via independent smoothing approach
2. Perform an fPCA analysis on $\hat{X}_1, \ldots, \hat{X}_n$

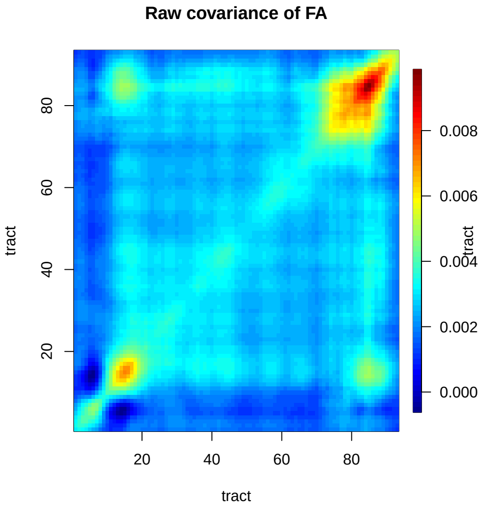
# FPCA of DTI Data (Smoothing)

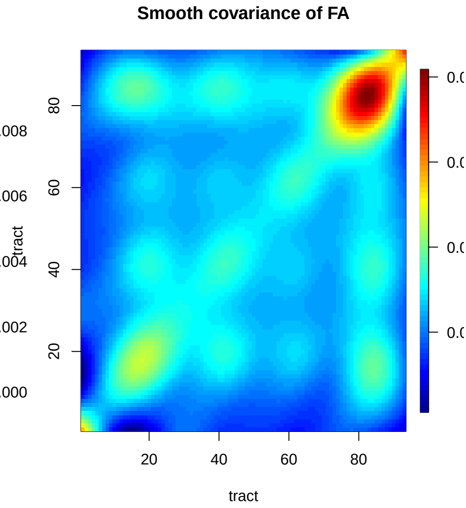
```
# Obtain Smooth Estimate of DTI data
s.dti <- array(0, dim(DTI.baseline$cca))
for(j in 1:nrow(DTI.baseline$cca)){
  fit <- gam(DTI.baseline$cca[j,] ~ s(tract), method = "REML")
  s.dti[j,] <- fit$fitted
}
```



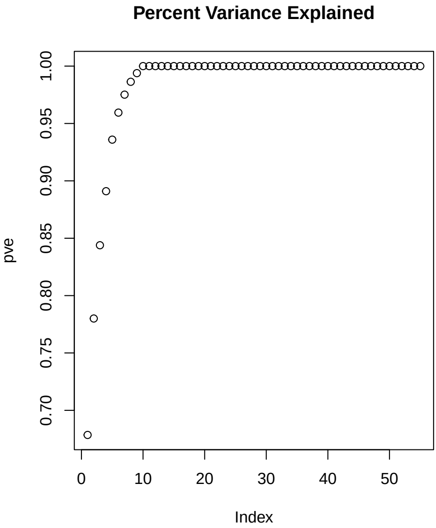
# fPCA of DTI Data (Covariance Estimation)

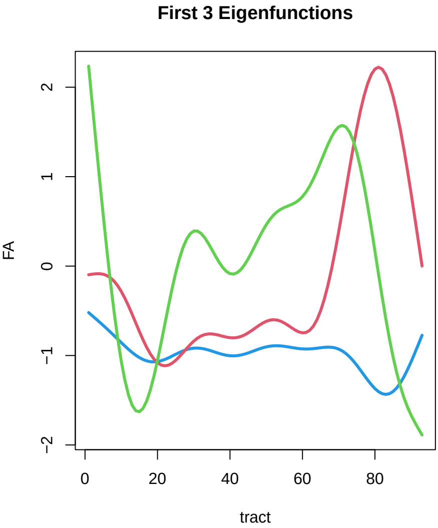
```
raw.cov <- cov(DTI.baseline$cca) # Raw Covariance
s.cov   <- cov(s.dti)            # Raw Covariance
```



# fPCA of DTI Data (Eigen-Analysis)

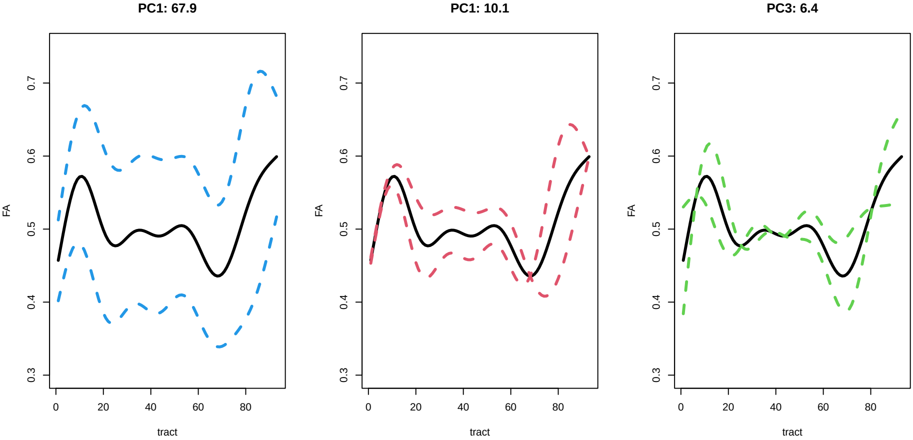


# fPCs as Components of Variation

- Sources of variance along fPCs are often visualized as

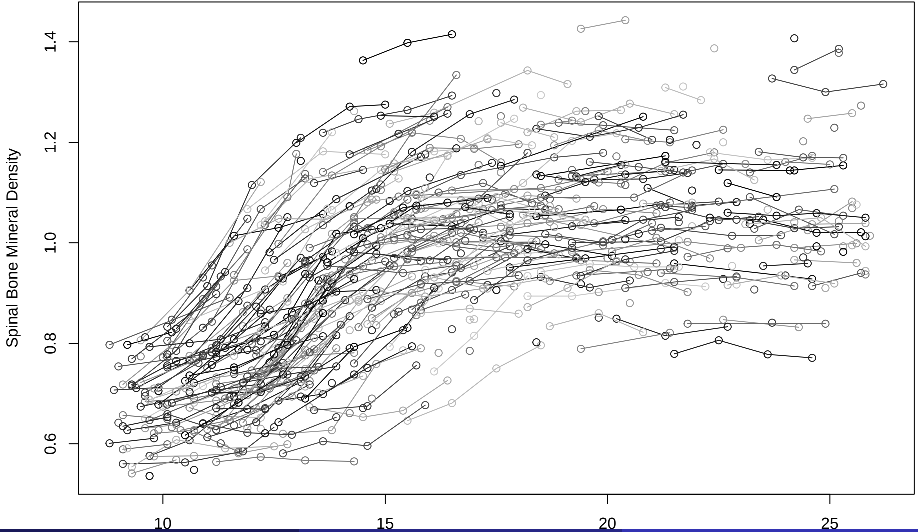
$$\hat{\mu}(t) \pm 2\sqrt{\hat{\lambda}_j}\hat{\phi}_j(t)$$



FPCA for sparse data.

# Data: Bone Mineral Density

- Relative spinal bone mineral density measurements on 261 North American adolescents.



# Estimation of FPCs

- Observed data $\{Y_i(t_{ij}),\ j = 1, \ldots, m_i\}_{i=1}^n$ iid
- Assume

$$Y_i(t_{ij}) = X_i(t_{ij}) + \epsilon_{ij};\ \ X_i \in L^2[T],\ \ \epsilon_{ij} \sim N(0, \sigma_\epsilon^2(t))$$

**Dense Sampling**:

1. Obtain $\hat{X}_i$ via independent smoothing approach
2. Perform an fPCA analysis on $\hat{X}_1, \ldots, \hat{X}_n$

**Sparse Sampling**

1. Many samples only provide a partial view of the realization path of $X_i(t)$
2. Subject-level smoothing is often not appropriate to obtain a reasonable

# Estimation Procedure in Sparse Settings

- Assume

$$Y_i(t_{ij}) = X_i(t_{ij}) + \epsilon_i(t_{ij})$$

with $E\{X_i\} = \mu$, $Cov(X_i) = c$, and $\epsilon_i(t_{ij}) \sim WN(\sigma_\epsilon^2(t))$

- This implies

$$E\{Y_i(t)\} = \mu(t)$$

and

$$Cov\{Y_i(s), Y_i(t)\} = c(s,t) + \sigma_\epsilon^2(t) I(t = s)$$

- Estimation of $\mu$ and $c$ relies on pooling data

# Estimation of the Mean in Multivariate Models

- Let $(Y_1, \ldots, Y_n) \in \mathbb{R}^k$ iid with mean $\mu$ and covariance $\Sigma$
- Estimation of $\mu$ and $\Sigma$ may be based on the minimizing

$$\Delta(\mu, \Sigma) = \sum_{i=1}^{n} (Y_i - \mu)^T \Sigma^{-1} (Y_i - \mu)$$

- **Lemma**: If $\Sigma \geq 0$, $\hat{\mu} = \arg\max_{\mu} \Delta(\mu, \Sigma) = \bar{X}$
- The objective $\Delta$, may be obtained as the negative log-likelihood assuming $X_i \sim N(\mu, \Sigma)$
- Equivalently, a GEE approach would base the estimation of $\mu$ on unbiased estimating equations of the form:

$$G_n(\hat{\mu}) = \sum_{i=1}^{n} \Sigma^{-1} (Y_i - \hat{\mu}) := 0$$

The GEE view allows us to show $\hat{\mu}$ is a CAN estimator

# Estimation of $\mu(t)$

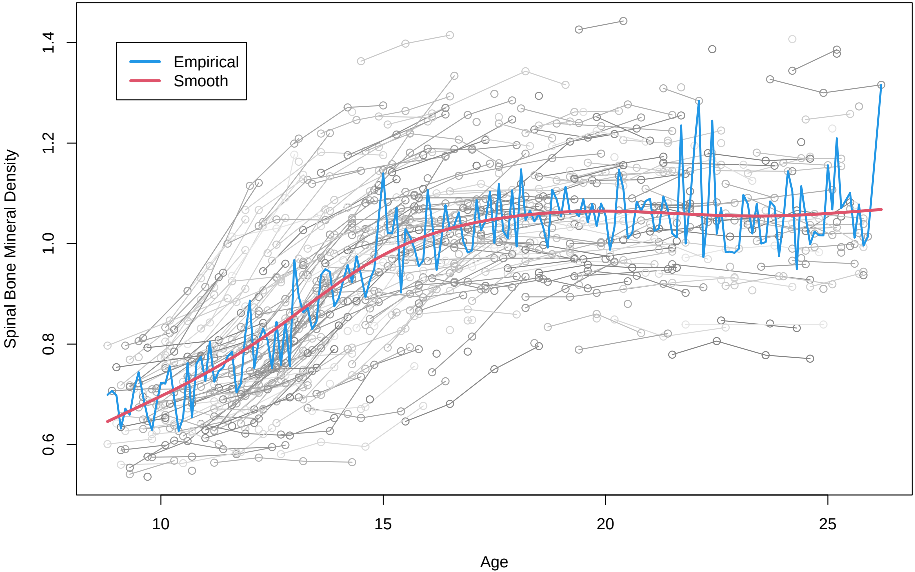
- In sparse settings $\bar{Y}$ is likely to be unsmooth
- Improved estimation is achieved by smoothing the data $\{(t_{ij}, Y_{ij}), i_1^n, j_1^{m_i}\}$
- Borrowing from the finite dimensional case PLS estimation ignores $\Sigma$, and minimizes

$$L(\mu) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\{Y_i(t_{ij}) - \mu_i(t_{ij})\}^2 + \lambda_0 \sum_{i=1}^{n}\sum_{j=1}^{m_i} \mu''(t_{ij})^2$$

- Representing $\mu$ with a set of basis functions, s.t. $\mu(t) = S(t)^T\beta$, the PLS criterion minimizes

$$L(\mu) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\{Y_i(t_{ij}) - S(t_{ij})^T\beta\}^2 + \lambda_0 \sum_{i=1}^{n}\sum_{j=1}^{m_i} \beta^T D \beta^T$$

# Smooth Estimate of $\mu$ (REML)



# Estimation of $c(s, t)$

- We assume $Cov\{Y(s), Y(t)\} = c(s,t) + \sigma_\epsilon^2(t) I(s = t)$
- The noise component defines a discontinuity on the diagonal, so some care is needed when smoothing

**Smoothing the Empirical Covariance**

- For any sampling pair $(t_{ij}, t_{ij'})$, $i = 1, \ldots, n$, evaluate the raw covariance

$$G_{ijj'} = \{X_i(t_{ij}) - \hat{\mu}(t_{ij})\}\{X_i(t_{ij'} - \hat{\mu}(t_{ij'})\}$$

- Smooth the empirical covariance $\{G_{ijj'},\ t_{ij},\ t_{ij'}\}_{i=1}^n$

1. Remove the diagonal temrs $G_{ijj}$
2. Define a large vector $G = \{G_{ijj'},\ j \neq j'\}_{i=1}^n$ and a matrix summarizing the evaluation points in two dimensions $T = \{(t_{ij}, t_{ij'}),\ j \neq j'\}_{i=1}^n$
3. Construct a bivariate smoother for $G = \Sigma + E$

# Bivariate Smoothing

Consider the model $Y(x,z) = \mu(x,z) + \epsilon(x,z)$.

- **Def.** The function $\mu(x,z)$ is represented using a **tensor product** basis of two variables, if given two sets of basis functions $B^{(1)}(x) \in \mathbb{R}^{p_x}$ and $B^{(2)}(z) \in \mathbb{R}^{p_z}$, and a matrix $\theta \in \mathbb{R}^{p_x \times p_z}$ of coefficients, we say

$$\mu(x,z) = [B^{(1)}(x)]^T \theta \; B^{(2)}(z)$$

**Bivariate Smoothing**

- Let $Y \in \mathbb{R}^{n1 \times n2}$, be a matrix of data observed on a set of $n_1$ values of $x$ and $n_2$ values of $z$.
- Let $B^{(1)} \in \mathbb{R}^{n_1 \times p_x}$ and $B^{(2)} \in \mathbb{R}^{n_2 \times p_z}$ be two sets of basis functions
- Bivariate smoothing via tensor products is obtained minimizing

$$||Y - B\theta B||^2 + \lambda \theta^T P \, \theta$$

# Bivariate Smoothing

- Let $\tilde{Y} = \text{vec}(Y)$. Tensor product smoothing can be expressed as follows
- Let $S_j = B^{(j)} \left( [B^{(j)}]^T B + \lambda_j D^{(j)} \right)^{-1} [B^{(j)}]^T$, $j = 1, 2$
- Tensor product smoothing results in the linear smoother

$$\text{vec}(\hat{\boldsymbol{\mu}}) = (S_1 \otimes S_2)\tilde{Y}$$

- Typically, a bivariate smoother applied to $G$, results in a symmetric estimate of $\tilde{\Sigma}$, e.g. if $S_1 = S_2$,
- $\tilde{\Sigma}$ is not guaranteed to be p.d. $\rightarrow$ Expand $\tilde{\Sigma} = U\tilde{\Lambda}U'$, with $\tilde{\Lambda} = \text{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m)$, [$m$ large evaluation grid].
- Choose a finite truncation $k$ (e.g. PEV $> 95\%$), with $\tilde{\lambda}_k > 0$ so that

$$\hat{c} = U^{(k)}\text{diag}(\hat{\lambda}_1, \ldots, \hat{\lambda}_k)[U^{(k)}]^T$$

involving only $k$ eigenvectors and eigenvalues

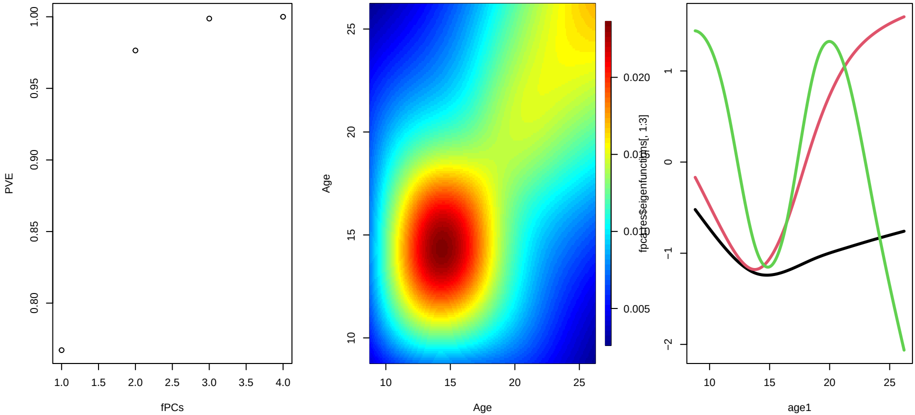
# Sparse fPCA

- Given $\hat{c}$ through smoothing and p.d. correction, estimate $\sigma_\epsilon^2(t)$
- One-dimensional smoothing $\{G_{ij},\ j = 1, \ldots, m_i\}_i^n$ leads to $\tilde{\sigma}_\epsilon^2(t)$
- The final estimator of $\sigma_\epsilon^2(t)$ takes the form

$$\hat{\sigma}_\epsilon^2(t) = \tilde{\sigma}_\epsilon^2(t) - \hat{c}(t, t)$$

- An eigenexpansion of $\hat{c}$ produces the eigenfunctions and eigenvalues $\left\{\hat{\lambda}_\ell, \hat{\phi}_\ell\right\}_{\ell=1}^k$
- Several details and alternatives in estimation are possible to arrive at fPCA estimates with similar large sample properties

# FPCA of Spinal Bone Mineral



# Estimation of fPC Scores

- Recall that through KL expansion we say

$$Y_i(t) = \mu(t) + \sum_{j=1}^{\infty} \nu_{ij}\phi_j(t) + \epsilon_i(t)$$

- In dense settings we use

$$\hat{\nu}_{ij} = \int (Y_i(t) - \hat{\mu}(t))\hat{\phi}_j(t)\, dt$$

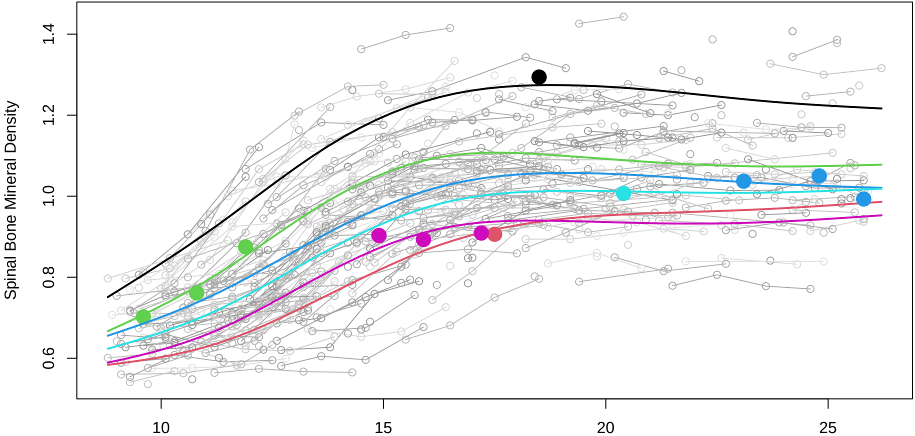
- If $Y_i(t)$ is sparsely observed using the integral estimator is no longer feasible
- Assuming Normality use BLUP

$$\tilde{\nu}_{ij} = \hat{\lambda}_j \hat{\phi}_j(\mathbf{t}_i)^T \{\hat{c}(\mathbf{t}_i, \mathbf{t}_i) + \mathrm{diag}(\hat{\sigma}^2_\epsilon(\mathbf{t}_i))\}^{-1}(Y_i - \hat{\mu}(\mathbf{t}_i))$$

# Reconstruction of the signal for $Y_i(t)$

At any time point $t$, the true signal for $Y_i(t)$ is obtained as

$$\hat{Y}_i(t) = \hat{\mu}(t) + \sum_{j=1}^{k} \hat{\nu}_{ij}\hat{\phi}_j(t)$$



# Theoretical Properties of Sparse fPCA

Under suitable regularity conditions

- The local-linear estimator of $\mu$

$$\sup_t |\hat{\mu}(t) - \mu(t)| = O_p(n^{-3/10})$$

- The local-linear estimator of $c$

$$\sup_{t,s} |\hat{c}(s,t) - c(s,t)| = O_p(n^{-1/10})$$

# Thank You!

Questions? Comments?