skfda.preprocessing.registration.mse_decomposition

skfda.preprocessing.registration.mse_decomposition(*original_fdata, registered_fdata, warping_function=None, *, eval_points=None*) [source]

Compute mean square error measures for amplitude and phase variation.

Once the registration has taken place, this function computes two mean squared error measures, one for amplitude variation, and the other for phase variation. It also computes a squared multiple correlation index of the amount of variation in the unregistered functions is due to phase.

Let $x_i(t), y_i(t)$ be the unregistered and registered functions respectively. The total mean square error measure (see [RGS09-8-5]) is defined as

$$\text{MSE}_{total} = \frac{1}{N}\sum_{i=1}^{N}\int[x_i(t)-\overline{x}(t)]^2 dt$$

We define the constant C_R as

$$C_R = 1 + \frac{\frac{1}{N}\sum_i^N \int[Dh_i(t)-\overline{Dh}(t)][y_i^2(t)-\overline{y^2}(t)]dt}{\frac{1}{N}\sum_i^N \int y_i^2(t)dt}$$

Whose structure is related to the covariation between the deformation functions $Dh_i(t)$ and the squared registered functions $y_i^2(t)$. When these two sets of functions are independents $C_R = 1$, as in the case of shift registration.

The measures of amplitude and phase mean square error are

$$\text{MSE}_{amp} = C_R \frac{1}{N}\sum_{i=1}^{N}\int\left[y_i(t)-\overline{y}(t)\right]^2 dt$$

$$\text{MSE}_{phase} = \int\left[C_R\overline{y}^2(t)-\overline{x}^2(t)\right]dt$$

It can be shown that

$$\text{MSE}_{total} = \text{MSE}_{amp} + \text{MSE}_{phase}$$

The squared multiple correlation index of the proportion of the total variation due to phase is defined as:

$$R^2 = \frac{MSE_{phase}}{MSE_{total}}$$

See [KR08-3] for a detailed explanation.

Parameters:

- **original_fdata** (`FData`) – Unregistered functions.
- **regfd** (`FData`) – Registered functions.
- **warping_function** (`FData`) – Warping functions.
- **eval_points** – (array_like, optional): Set of points where the functions are evaluated to obtain a discrete representation.

Returns: Tuple with amplitude mean square error MSE_{amp}, phase mean square error MSE_{phase}, squared correlation index R^2 and constant C_R.

Return type: `collections.namedtuple`

Raises: `ValueError` – If the curves do not have the same number of samples.

References

[KR08-3] Kneip, Alois & Ramsay, James. (2008). Quantifying amplitude and phase variation. In *Combining Registration and Fitting for Functional Models* (pp. 14-15). Journal of the American Statistical Association.

[RGS09-8-5] Ramsay J.O., Giles Hooker & Spencer Graves (2009). In *Functional Data Analysis with R and Matlab* (pp. 125-126). Springer.

Examples

```
>>> from skfda.datasets import make_multimodal_landmarks
>>> from skfda.datasets import make_multimodal_samples
>>> from skfda.preprocessing.registration import (landmark_registration_warping,
...                                               mse_decomposition)
```

We will create and register data.

```
>>> fd = make_multimodal_samples(n_samples=3, random_state=1)
>>> landmarks = make_multimodal_landmarks(n_samples=3, random_state=1)
>>> landmarks = landmarks.squeeze()
>>> warping = landmark_registration_warping(fd, landmarks)
>>> fd_registered = fd.compose(warping)
>>> mse_amp, mse_pha, rsq, cr = mse_decomposition(fd, fd_registered, warping)
```

Mean square error produced by the amplitude variation.

```
>>> f'{mse_amp:.6f}'
'0.000987'
```

In this example we can observe that the main part of the mean square error is due to the phase variation.

```
>>> f'{mse_pha:.6f}'
'0.115769'
```

Nearly 99% of the variation is due to phase.

```
>>> f'{rsq:.6f}'
'0.991549'
```