STATISTICAL LEARNING SYSTEMS

LECTURE 2: LINEAR REGRESSION AND BEYOND

Jacek Koronacki

Institute of Computer Science, Polish Academy of Sciences

Ph. D. Program 2013/2014

The project is co-financed by the European Union within the framework of European Social Fund

Principal Components Regression (PCR):

Use the principal components as the derived input columns $\mathbf{z}_m = \mathbf{X}\mathbf{v}_m$, and then regress $\mathbf{y}$ on $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_M$ for some $M \leqslant p$. Since the $\mathbf{z}_m$ are orthogonal, this regression is just the sum of univariate regressions:

$$\hat{\mathbf{y}}^{\mathrm{pcr}}_{(M)} = \bar{y}\mathbf{1} + \sum_{m=1}^{M} \hat{\theta}_m \mathbf{z}_m,$$

where $\hat{\theta}_m =< \mathbf{z}_m, \mathbf{y} > / < \mathbf{z}_m, \mathbf{z}_m >$. Since the $\mathbf{z}_m$ are each linear combinations of the original $\mathbf{x}_j$, we can express the above in terms of coefficients of the $\mathbf{x}_j$ (for $M = p$ we get the usual LS estimates):

$$\hat{\beta}^{\mathrm{pcr}}(M) = \sum_{m=1}^{M} \hat{\theta}_m \mathbf{v}_m.$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - derived input directions

Since principal components depend on the scaling of the inputs, we typically first standardize them (just as in ridge regression).

There is an obvious relationship between PCR and ridge regression: while the latter shrinks the coefficients of the principal components (the more it shrinks the smaller the corresponding eigenvalue is), the former discards the $p - M$ smallest eigenvalue components.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - derived input directions

Partial Least Squares (PLS):

- Standardize each $\mathbf{x}_j$ and set $\hat{\mathbf{y}}^{(0)} = \bar{y}\mathbf{1}$, and $\hat{\mathbf{x}}_j^{(0)} = \mathbf{x}_j$, $j = 1, \ldots, p$.
- For $m = 1, 2, \ldots, p$
  - $\mathbf{z}_m = \sum_{j=1}^{p} \hat{\phi}_{mj}\mathbf{x}_j^{(m-1)}$, where $\hat{\phi}_{mj} = < \mathbf{x}_j^{(m-1)}, \mathbf{y} >$.
  - $\hat{\theta}_m = < \mathbf{z}_m, \mathbf{y} > / < \mathbf{z}_m, \mathbf{z}_m >$.
  - $\hat{\mathbf{y}}^{(m)} = \hat{\mathbf{y}}(m-1) + \hat{\theta}_m\mathbf{z}_m$.
  - $\mathbf{x}_j^{(m)} = \mathbf{x}_j^{(m-1)} - [< \mathbf{z}_m, \mathbf{x}_j^{(m-1)} > / < \mathbf{z}_m, \mathbf{z}_m >]\mathbf{z}_m$, $j = 1, 2, \ldots, p$.
- Output the fitted vectors $\{\hat{\mathbf{y}}^{(m)}\}_1^p$. Since the $\{\mathbf{z}_\ell\}_1^m$ are linear in the original $\mathbf{x}_j$, so is $\hat{\mathbf{y}}^{(m)} = \mathbf{X}\hat{\beta}^{\mathrm{pls}}(m)$. (These linear coefficients can be recovered from the sequence of PLS transformations.)

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - derived input directions

It can be shown that PLS seeks directions that have high variance and have high correlation with the response, in contrast to PCR which keys only on high variance. Indeed, the mth principal component direction $\mathbf{v}_m$ solves:

$$\max_\alpha \mathrm{Var}(\mathbf{X}\alpha)$$

subject to $||\alpha|| = 1$, $\alpha' \mathbf{S}\mathbf{v}_\ell = 0$, $\ell = 1, \ldots, m-1$, while the mth PLS direction $\hat{\phi}_m$ solves:

$$\max_\alpha \mathrm{Corr}^2(\mathbf{y}, \mathbf{X}\alpha)\mathrm{Var}(\mathbf{X}\alpha)$$

subject to $||\alpha|| = 1$, $\alpha' \mathbf{S}\hat{\phi}_\ell = 0$, $\ell = 1, \ldots, m-1$. (However, further analysis has revealed that the variance aspect tends to dominate.)

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - a small coda for now

We have dwelt for quite some time on the lasso and LAR, the latter providing the whole solution path for the former. The lasso problem can be written as

$$\text{minimize}_{\boldsymbol{\beta}\in\mathrm{R}^{\mathrm{p}}}\frac{1}{2}||\mathbf{y}-\mathbf{X}\boldsymbol{\beta}||_2^2+\lambda||\mathbf{D}\boldsymbol{\beta}||_1,$$

where $\mathbf{D}=\mathbf{I}$. In the above, we have included a penalty matrix $\mathbf{D}$ to hint to the Generalized Lasso - see R.J Tibshirani and J. Taylor, Ann. Statist. (2011), Vol. 39, No. 3, 1335-1371. We shall come back to the generalized lasso in another context in due time.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - Generalized Linear Models (GLM)

While linear regression models are truly versatile and general, they can hardly be recommended when, say, the response variable is binary, let alone when it has a Poisson distribution. Also, when we know (up to unknown parameters) its probability distribution, replacing the LS method by maximum likelihood estimation is most often recommended (not only when the responses are discrete but also when their distributions are continuous, in particular, when they are skewed).

It is here where Generalized Linear Models (GLM) can prove to be of great help. When defining them, we shall shall show simultaneously how they generalize the linear model with normal errors.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - GLM

- The random component: $\mathbf{Y}$ (the vector of observations of the response variable) has i.i.d. components and has mean $E(\mathbf{Y}) = \boldsymbol{\mu}$; in the case of the linear model, each

$$y_i \sim N(\mu_i, \sigma^2).$$

- The systematic component: covariates (explanatory variables or predictors) $x_0 = 1, x_1, x_2, \ldots x_p$ produce a linear predictor η given by

$$\eta = \mathbf{x}'\boldsymbol{\beta},$$

  where $\boldsymbol{\beta}$ is a vector of parameters.
- The link function $\ell(\cdot)$ describes how the mean response, $E(y) = \mu$, is linked to the systematic component:

$$\eta = \ell(\mu);$$

  in the case of the linear model, $\eta = \mu$.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - GLM

Another particular case of a GLM is that of logistic regression which, for a single observation, assumes the form (logistic regression will be discussed in greater detail within a section on linear discriminant analysis):

$$\eta = \alpha + \beta^T \mathbf{x} = \log \frac{p(2|\mathbf{x})}{1 - p(2|\mathbf{x})} \equiv \text{logit}(p(2|\mathbf{x});$$

here, Y is a proportion of successes (getting class 2) in n Bernoulli trials conditioned upon $\mathbf{x}$, with

$$E(y|\mathbf{x}) = p(y = 2|\mathbf{x});$$

accordingly, the link function $\ell(\cdot)$ is given by

$$\eta = \ell(p(2|\mathbf{x})) = \text{logit}(p(2|\mathbf{x}).$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

More often than not, for count regression, Poisson regression does the job.

It is easy to show that for Poisson regression the link function is given by the natural logarithm.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - GLM

A beautiful theory of GLM exists which covers the case when response variables $\mathbf{Y}$ are from the exponential family of distributions:

$$h(y;\nu,\phi) = \exp\left[\frac{y\nu - b(\nu)}{a(\phi)} + c(y,\phi)\right],$$

where $a(\cdot)$, $b(\cdot)$ i $c(\cdot)$ are known functions. The ν is called the canonical parameter and represents location while ϕ is called the dispersion parameter and represents the scale.

It is easily seen that, e.g., normal, binomial, Poisson and gamma distributions belong to the exponential family of distributions.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

We have:

$$Ey = \mu = b'(\nu)$$
$$\mathrm{Var}y = a(\phi)b''(\nu)$$

The mean is a function of ν only, while the variance is a product of functions of the location and scale; $b''(\nu)$ is called the variance function and describes how the variance relates to the mean.

Remark: For a normal distribution, $a(\phi) = \phi = \sigma^2$; for binomial and Poisson distributions, $a(\phi) = \phi = 1$.

In general, we assume:

$$a(\phi) = \phi/w,$$

where w is a known weight that varies between observations.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

The canonical link has $\ell(\cdot)$ such that $\eta = \ell(\mu) = \nu$, the canonical parameter of the exponential family distribution. This means that $\ell(b'(\nu)) = \nu$. From now on we assume that the GLMs under study are in canonical form.

The canonical link for the normal distribution is the identity function, $\eta = \mu$. The canonical link for the binomial distribution (and the response equal to the proportion of successes, not the number of successes, e.i., when $\mu = p$, where p is the probability of success), the canonical link has the logit link function, $\eta = \log(p/(1-p))$. For the Poisson distribution, the canonical link has the natural logarithm as its link function.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - GLM

The parameters β of a GLM are estimated using maximum likelihood, actually by applying the Newton-Raphson method which can be shown to be equivalent to Iteratively Reweighted Least Squares (IRWLS).

Let us omit the details and note only that in this approach we rely on a linearized form of the link function applied to the data:

$$\ell(y) \approx \ell(\mu) + (y - \mu)\ell'(\mu) = \eta + (y - \mu)\frac{d\eta}{d\mu} \equiv z,$$

and we regress z on the x_j using IRWLS with weights w_i for successive observations given by (from now on we assume that $a(\phi) = 1$)

$$w_i^{-1} = (\frac{d\eta}{d\mu}|_{\mu=\hat{\mu}_i})^2 b''(\hat{\nu}_i),$$

where $\hat{\mu}_i$ and $\hat{\nu}_i$ are current (i.e. corresponding to the current iteration of the algorithm) approximations to μ_i and ν_i, $i = 1, \ldots, n$ ($\nu_i = \eta_i$).

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Accordingly, the basis for assessing the quality of fit of a particular GLM ω is provided by the model deviance

$$dev_\omega = 2 \log \frac{L_{\omega_{full}}}{L_\omega} = 2(\log L_{\omega_{full}} - \log L_\omega),$$

where L_ω is the likelihood for the model under consideration, ω, and $L_{\omega_{full}}$ is the likelihood for the full (or saturated) model, ω_{full}. It follows that dev_ω is always nonnegative.

Remark: For the Gaussian regression model, deviance is, up to a constant, equal to $\mathrm{RSS} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Under some natural assumptions, we can compare a larger GLM, Ω, from an exponential family of distributions with a smaller nested GLM, ω, since, under the hypothesis that the smaller model is adequate,

$$dev_\Omega - dev_\omega$$

is asymptotically χ^2 distributed with with degrees of freedom equal to the difference in the number of parameters in the two models.

Remark: One may construct a z-statistic for, say, β_p,

$$\frac{\hat{\beta}_p}{\mathrm{SE}(\hat{\beta}_\mathrm{p})}$$

to use the Wald test to check whether β_p in the model is zero (against the alternative that it is not). However, the difference of deviances test is preferred to the Wald test, and for good reasons.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

One can use the following counterpart of the coefficient of determination R^2

$$1 - \frac{dev_\omega}{dev_{\omega_0}}$$

where dev_{ω_0} is the deviance of a GLM $\omega_0 : y \sim 1$, which includes only a constant as the sole explanatory variable, as an obvious measure of fit for the given GLM ω.

Remark: For the Akaike criterion, we have $AIC = dev_\omega + 2(p+1)$.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - GLM diagnostics

When it comes to finding possible influential observations and/or outliers for a GLM ω, one has to start with a proper definition of residuals as well as to turn to the concept of leverages.

We shall confine ourselves to introducing deviance residuals, $r_{dev,i}$, $i = 1, 2, \ldots, n$. Note that they should satisfy the relationship

$$\sum r_{dev,i}^2 = dev_\omega;$$

hence, they are given the form

$$r_{dev,i} = sign(y - \hat{\mu})\sqrt{dev_{\omega,i}}.$$

Of course, each $dev_{\omega,i}$ is a component of dev_ω which corresponds to the ith observation.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

The residuals may be studentized as follows:

$$\frac{r_{dev,i}}{\sqrt{1-h_{ii}}},$$

where h_{ii} are the diagonal elements, the so-called leverages, of the hat matrix,

$$\mathbf{H} = \mathbf{W}^{1/2}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{1/2},$$

or one can compute jacknife residuals. Both can help find outliers, although the former may fail when an observation is influential. The leverages alone represent the potential of the observation to influence the fit.

Another, and a better, way to find influential observations is to examine the Cook statistic:

$$(\hat{\beta}_{(i)} - \hat{\beta})'(\mathbf{X}'\mathbf{W}\mathbf{X})(\hat{\beta}_{(i)} - \hat{\beta})/p.$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - GLM diagnostics

A way to look for unusual observations is to make half-normal plots of the residuals (e.g. jacknife residuals) and of the leverages. Of course, the plots of residuals against fitted values (usually the linear predictors $\hat{\eta}$) can be of help too.

Partial residual plots of the type

$$z - \hat{\eta} + \hat{\beta}_j x_j \quad \text{versus} \quad x_j$$

as well as such plots as the plot of linearized response z versus the linear predictor can help check the assumptions of the model. (Partial residual plots play the same role as in the case of linear models, where $\hat{\varepsilon} + \beta_j x_j$ is plotted against x_j; i.e., they show the relationship between the linearized response and a given predictor after having taken into account the effect of other predictors.)

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

It is clear that the ANOVA (analysis of variance) and ANCOVA (analysis of covariance) models are examples of linear regression models with design matrices $\mathbf{X}$ of a specific (and simple) structure. The simplest of them, i.e. the one-way ANOVA model, can be written as

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij},$$

where $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n_i$, Y_{ij} is the jth observation with fixed effect α_i, i.e., with factor α at its ith level, μ is the overall mean, and the ε_{ij} are i.i.d. errors which are assumed to be normally distributed, $N(0, \sigma^2)$, with unknown σ. For each fixed i, we say that the observations Y_{ij}, $j = 1, 2, \ldots, n_i$, form a group. Fixed effects α_i are constant and unknown. In order to make the model well-defined, we assume also that

$$\alpha_1 + \alpha_2 + \cdots + \alpha_k = 0.$$

PhD STUDIES
KAPITAŁ LUDZKI NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA EUROPEJSKI FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Sometimes, the effects α_i should not be considered fixed, i.e., the same over a group of observations, but random. In such a (one-way) random effects model we usually assume that the ε_{ij} and the effects α_i are uncorrelated, $\mathrm{Var}(\alpha_i) = \sigma_\alpha^2 \geqslant 0$ and $Var(\varepsilon_{ij}) = \sigma_\varepsilon^2 > 0$ for all i, j.

The unknown σ_α^2 and σ_ε^2 are called variance components. Notice that correlation between observations at the same level, the so-called intraclass correlation coefficient, is

$$\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}.$$

The question which should be asked is whether our claim that the effects α_i are random is correct. To put it otherwise, we should check if there is enough evidence to reject the following hypothesis:

$$H_0: \quad \sigma_\alpha^2 = 0.$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - mixed effect models

More often, we are faced with the situation where both fixed effects and random effects should be included into the model. In this way, we arrive at a mixed effects model, whose simplest version is that of a two-way mixed effects model with no interactions,

$$Y_{ijm} = \mu + \alpha_i + \beta_j + \varepsilon_{ijm},$$

where $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, \ell$, $m = 1, \ldots, n_{ij}$, μ is the overall mean, the α_i are fixed effects, the β_j are i.i.d normally distributed random effects, $N(0, \sigma_\beta^2)$, the errors ε_{ijm} are i.i.d. normally distributed errors, $N(0, \sigma_\varepsilon^2)$, and the random effects and errors are mutually independent.

It is easy to see that the intraclass correlation coefficient is

$$\frac{\sigma_\beta^2}{\sigma_\beta^2 + \sigma_\varepsilon^2}.$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - mixed effect models

More generally, the mixed effects model can be given the following form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{X}$ is an $n \times p$ model matrix, $\boldsymbol{\alpha}$ is a vector of p fixed parameters, $\mathbf{Z}$ is an $n \times q$ matrix associated with a vector $\boldsymbol{\beta}$ of q random effects and $\boldsymbol{\varepsilon}$ is vector of random errors.

We usually assume that $\boldsymbol{\beta} \sim N(0, \sigma^2\mathbf{D})$, $\boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I})$, the two vectors are independent, and hence $Y \sim N(\mathbf{X}\boldsymbol{\alpha}, \sigma^2(\mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{I}))$.

The parameters to be estimated are: $\boldsymbol{\alpha}$, σ^2 and the variance components $\mathbf{D}$.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund

Linear regression and beyond - mixed effect models

Maximum likelihood estimation (MLE, the abbrevation to be used for ML estimators too) of the unknown parameters, even for a one-way random effects model, is in fact not straightforward, to say the least (for reasons to be mentioned during the lecture).

Let us only mention here that one problem is with possible large bias of the MLE of the variance components associated with factors which have a small number of levels.

A way to get round this problem is to use restricted maximum likelihood (REML) estimators. The idea behind REML will be presented during the lecture.

Remark: Concluding, let us mention the ease with which the mixed effects models can be used to describe multilevel (hierarchical) models with nested effects.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
The project is co-financed by the European Union within the framework of European Social Fund