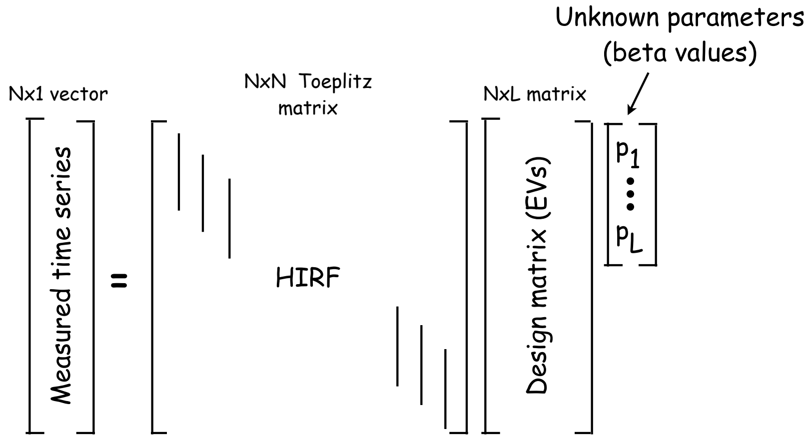
Modeling the fMRI time series



N: number of time points in the time series.
L: number of regressors in the design matrix.

Design matrix

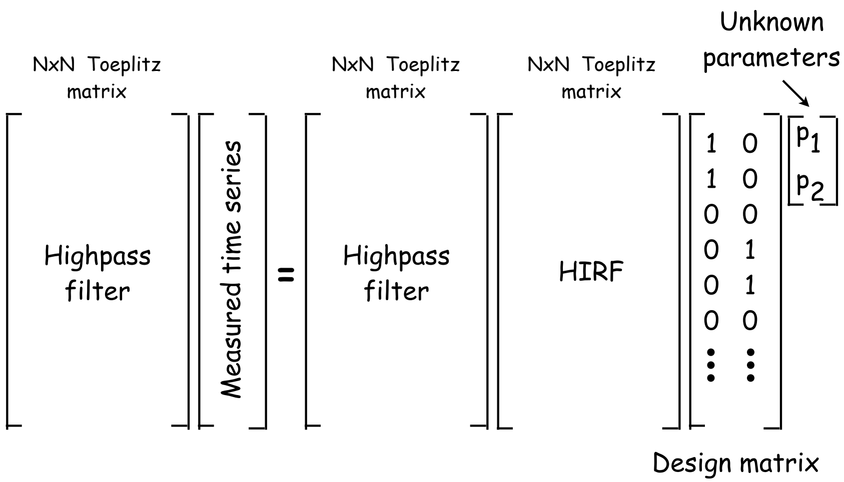
p_1 is amplitude of modulation of underlying neural activity

$$\begin{array}{cccc} \text{Block alternation (stimulus on/off)} & \text{Mean} & \text{Drift} & \\ \end{array}$$

$$\begin{bmatrix} 1 & 1 & -1 & \\ 1 & 1 & & \\ 1 & 1 & & \\ 0 & 1 & \vdots & \text{Other nuisance variables} \\ 0 & 1 & & \\ 0 & 1 & & \\ \vdots & \vdots & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \end{bmatrix}$$

Convolve only the relevant EVs (1st col in this example) with the HIRF, not the nuisance EVs.

Detrend during pre-processing

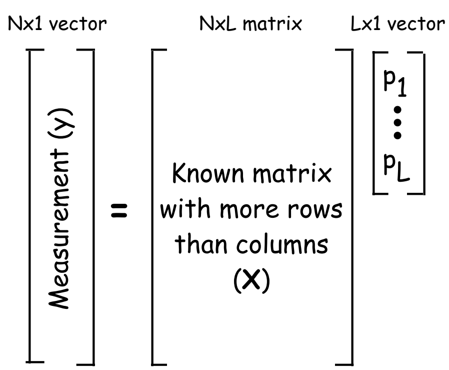


General linear model

Solve: $y = \mathbf{X}\, p$

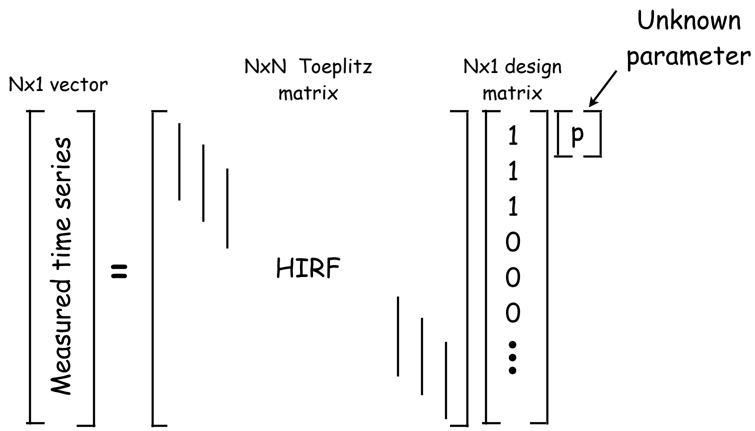
Answer: $p_{opt} = \mathbf{X}^{\#} y$

where p_{opt} are parameter estimates and # means pseudo-inverse.

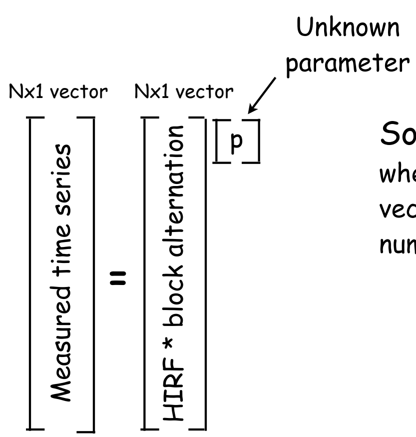


N: number of time points in the time series.
L: number of regressors in the design matrix.

Simple (one parameter) example



One parameter example (cont)



Solve: $\vec{y} = x\,\vec{p}$
where y and x are both vectors and p is a scalar number.

Least-squares regression

Find p to make x_n p as close as possible to y_n for all n. That is, choose p to minimize:

$$\min_p \sum_{n=1}^{N} (y_n - p x_n)^2$$

Or, in vector notation:

$$\min_p \|\vec{y} - p\vec{x}\|^2$$

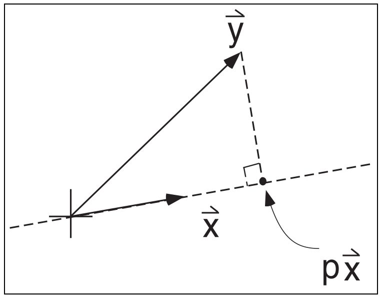
Solution 1 (using calculus). Take the derivative of the above expression, set it equal to zero, and solve for p:

$$p_{\text{opt}} = \frac{\vec{y}^T \vec{x}}{\vec{x}^T \vec{x}}.$$

Least-squares regression (cont)

Solution 2 (using geometry). Find the scale factor p such that the scaled vector p x is as close as possible (in Euclidean distance) to y. Geometrically, we know that the scaled vector should be the projection of y onto the line in the direction of x:

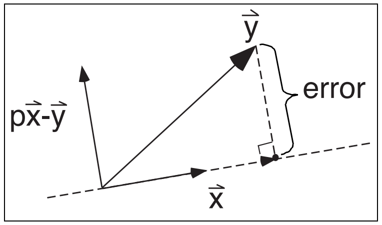
$$p\vec{x} = (\vec{y} \cdot \hat{x})\hat{x} = \frac{(\vec{y} \cdot \vec{x})}{\|\vec{x}\|^2}\vec{x}$$



Least-squares regression (cont)

Solution 3 (using the orthogonality principle). The error vector for the best p is perpendicular to x:

$$\vec{x} \cdot (p\vec{x} - \vec{y}) = 0.$$



Least-squares regression (multiple parameters)

Find best p for satisfying: $\vec{y} = \mathbf{X}\,\vec{p}$

Answer: $\vec{p}_{opt} = \mathbf{X}^{\#}\,\vec{y}$

where p_{opt} are parameter estimates and # is pseudo-inverse:

$$\mathbf{X}^{\#} = (\mathbf{X}^T \mathbf{X})^{-1}\,\mathbf{X}^T$$

One way to see this is as a generalization of the orthogonality principle. The error vector should be perpendicular to all of the basis vectors (columns of X):

$$\mathbf{X}^T\,(\vec{y} - \mathbf{X}\,\vec{p}) = 0$$

Solving for p gives the above expression.

Multiple regression (cont)

Find best p for satisfying: $\vec{y} = X\,\vec{p}$

$$\vec{p}_{opt} = X^{\#}\,\vec{y}$$

$$X^{\#} = (X^T X)^{-1} X^T$$

Another way to see this is algebraically:

$$(X^T X)^{-1} X^T \vec{y} = (X^T X)^{-1} X^T X\,\vec{p}$$

$$(X^T X)^{-1} X^T y = p_{opt}$$

Identity matrix

Multiple regression (cont)

Find best p for satisfying: $\vec{y} = X\,\vec{p}$

$$\vec{p}_{opt} = X^{\#}\,\vec{y}$$

$$X^{\#} = (X^T X)^{-1} X^T$$

Matlab code:

```
p = pinv(X)*y;
```

or

```
p = X\y;
```

Multiple regression (cont)

Find best p for satisfying: $\vec{y} = X\,\vec{p}$

$$\vec{p}_{opt} = X^{\#}\,\vec{y}$$

$$X^{\#} = (X^T X)^{-1} X^T$$

- Is $X^T X$ always invertible? If not, why not?
- What is the interpretation for the values corresponding to each element of $\vec{p}_{opt}$? Is the meaning of each value independent of the other elements?

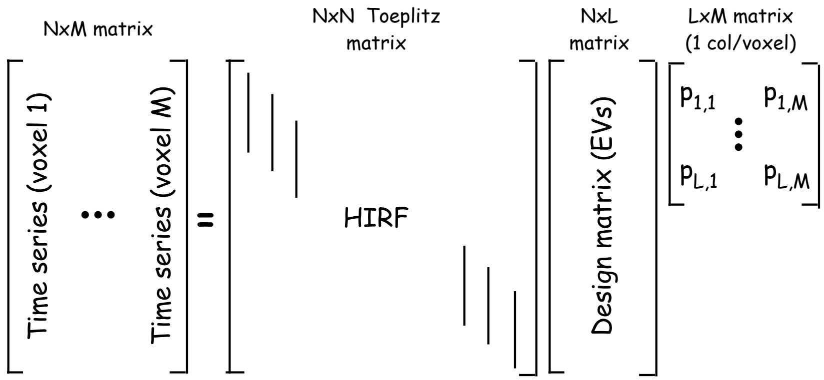
Trivial example

$$\begin{array}{ccccc} y & = & X & & p \\ \begin{bmatrix} 3 \\ 2 \\ .1 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} & & \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \end{array} \quad \Leftrightarrow \quad \begin{aligned} 3 &= 1\,p_1 + 0\,p_2 \\ 2 &= 0\,p_1 + 1\,p_2 \\ .1 &= 0\,p_1 + 0\,p_2 \end{aligned}$$

$$X^T X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (X^T X)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$(X^T X)^{-1} X^T y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ .1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

All the voxels at once

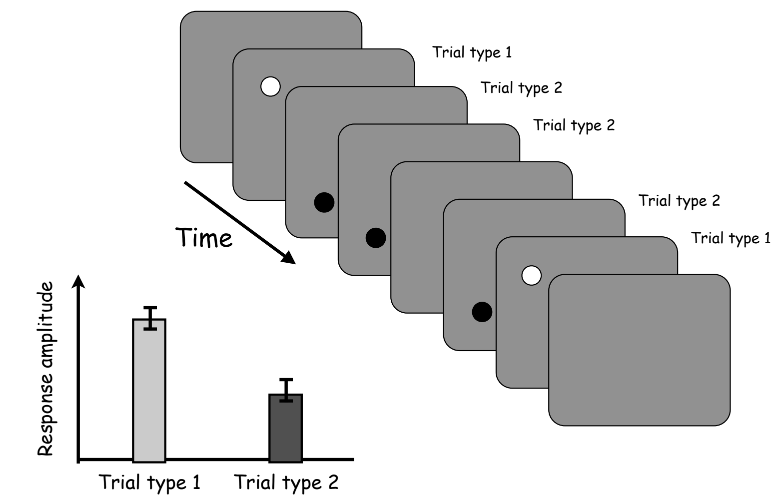


N: number of time points in the time series.
M: number of voxels.
L: number of regressors in the design matrix.

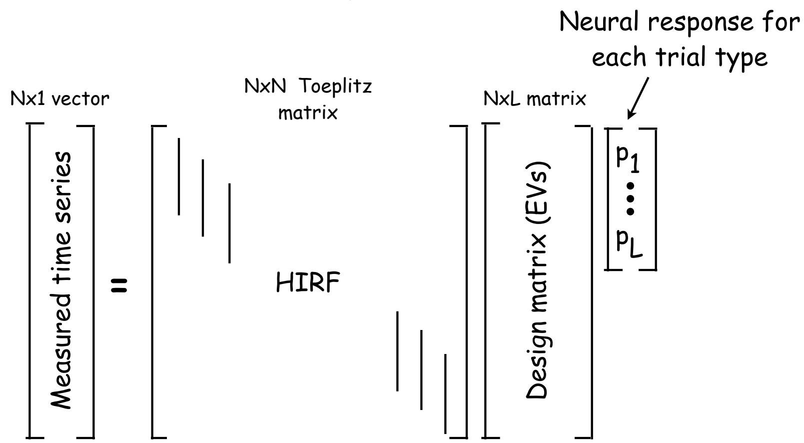
$$Y = X P$$

$$P_{opt} = X^{\#} Y$$

Event-related fMRI experiment

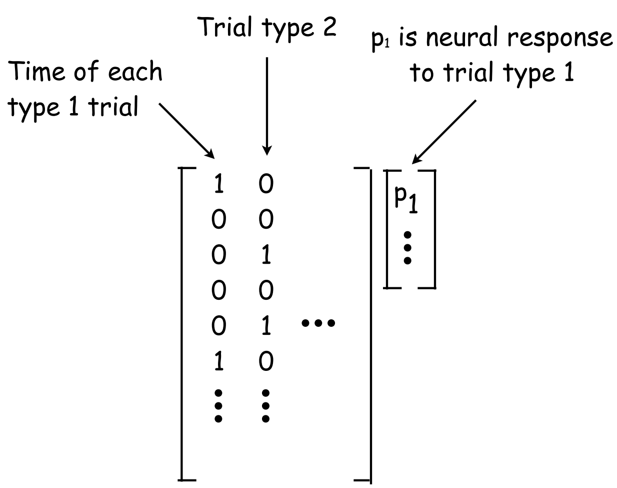


Event-related analysis

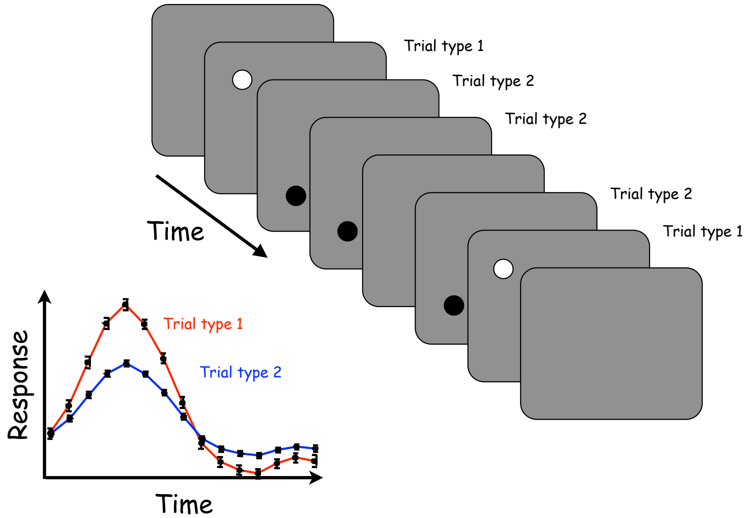


N: number of time points in the time series.
L: number of trial types.

Event-related design matrix



"Deconvolution" analysis



Deconvolution design matrix

$$\begin{array}{cccc}
\text{Nx1 vector} & & \begin{array}{c}\text{Nx2L matrix}\\ \text{Trial type 1} \qquad\qquad \text{Trial type 2}\end{array} & \\
\begin{bmatrix}\text{Measured time series}\end{bmatrix} & = & \left[\begin{array}{cccccc:cccccc}
1 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 0 & \\
0 & 1 & 0 & 0 & 0 & & 1 & 0 & 0 & 0 & 0 & \\
0 & 0 & 1 & 0 & 0 & & 0 & 1 & 0 & 0 & 0 & \\
1 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots \\
0 & 1 & 0 & 0 & 1 & & 0 & 0 & 0 & 1 & 0 & \\
0 & 0 & 1 & 0 & 0 & & 0 & 0 & 0 & 0 & 1 & \\
0 & 0 & 0 & 1 & 0 & & 0 & 0 & 0 & 0 & 0 & \\
 & & \vdots & & & & & & \vdots & & &
\end{array}\right] & \begin{bmatrix} p_1 \\ \vdots \\ p_{2L} \end{bmatrix}
\end{array}$$

N: number of time points in the time series.
L: number of time points in the HRF for each of 2 trial types.

Statistics

1. How well does the model fit the data?
2. What are the confidence intervals/error bars on the parameter estimates?
3. Are the parameter estimates different from zero? Different from each other?
4. Which of the regressors contribute to fitting the data?