ROI ANALYSIS OF PHARMAFMRI DATA: AN ADAPTIVE APPROACH FOR GLOBAL TESTING

Giorgos Minas[♣], John A.D. Aston[♢], Thomas E. Nichols[♡] and Nigel Stallard[♠]

[♣] Department of Statistics and Warwick Centre of Analytical Sciences,
[♢] CRISM, Department of Statistics,
[♡] Department of Statistics and Warwick Manufacturing Group,
[♠] Division of Health Sciences, Warwick Medical School,

University of Warwick, UK

46th Scientific Meeting Of The Italian Statistical Society,
Sapienza University of Rome - Faculty Of Economics
June 20-22, 2012

FUNCTIONAL MAGNETIC RESONANCE IMAGING

- fMRI: tool for studying **brain activity** using the BOLD contrast
- BOLD contrast: characterize relative local blood flow changes accompanying brain activity
- Data: **time-series of 3D brain images** recording the BOLD contrast
- Temporal resolution $\approx$ 3 seconds (few hundreds time-points)
- Spatial resolution $\approx$ 3 mm (tens of thousands locations)
- Use: Brain Mapping, Psychology, Marketing, Criminology, . . . , and *Clinical trials (CT): controlled experiment to compare the effects of different medical treatments on human subjects*
- **PharmafMRI**: Use of fMRI endpoints in clinical trials:
  - not standard clinical method but great potential
  - Application: Schizophrenia, Alzheimer, Addiction, Pain treatment, . . .

Background image source: http://fact0ry.blogspot.com

Region of Interest (ROI) analysis

Restrict data analysis to selected Regions of Interest (ROI). ***Why?***

- Easier to explore data
  - ▶ formulation of regional hypotheses
    - → more suitable for CTs

- To test regional hypotheses answering questions such as ...

"**does my drug change brain activity *in any* of regions A,B or C?**"

  - ▶ Drastic reduction of data dimensionality
    - → statistical power increased

ROI SUMMARY MEASURES

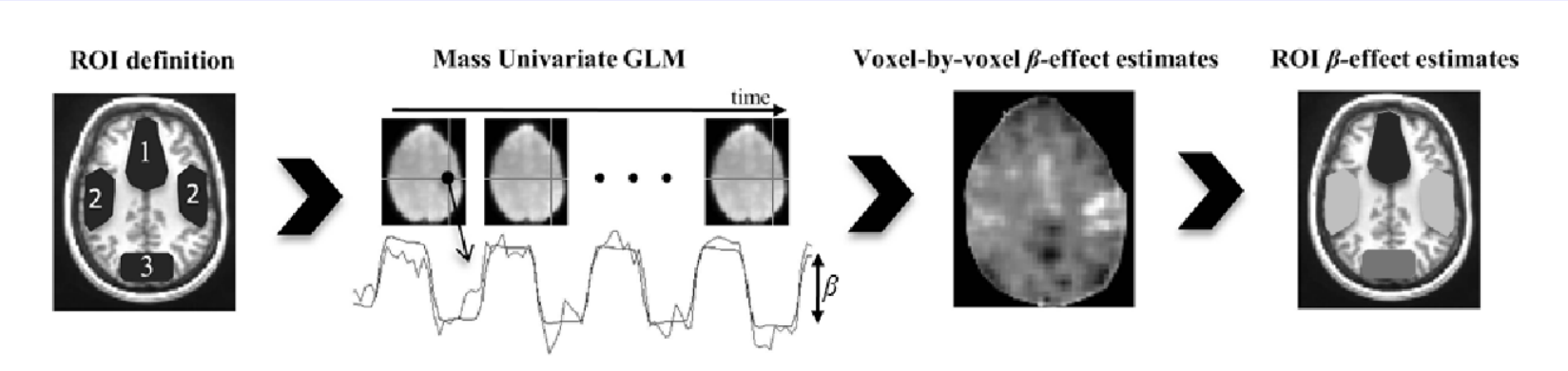


FIGURE: The steps for deriving ROI measures of treatment effect

To compute summary measure of the treatment effect in each ROI [1],

1. define exact ROI location (based on brain anatomy or function)
2. apply a suitable model (typically GLM) on each location time-series
3. extract the estimate of the treatment effect, $\hat{\beta}_k$, in voxel k
4. average $\hat{\beta}_k$'s within each ROI

TESTING FOR GLOBAL TREATMENT EFFECTS

- Responses:

$$Y_i = (Y_{i1}, Y_{i2}, ..., Y_{iK})^T,$$
$$E(Y_i) = \mu, \quad Var(Y_i) = \Sigma, \quad i = 1, \ldots, n, ind.$$

- Test the multivariate (MV) null hypothesis

$$\mathbf{H_0}: \ \mu = (\mu_1, ..., \mu_K)^T = (0, 0, ..., 0)^T = \underline{0}$$

BONFERRONI-TYPE METHODS multiple univariate tests controlling FWER

- conservative for high correlations-large K [2]

HOTELLING'S T^2 TEST likelihood ratio test for MV normal responses

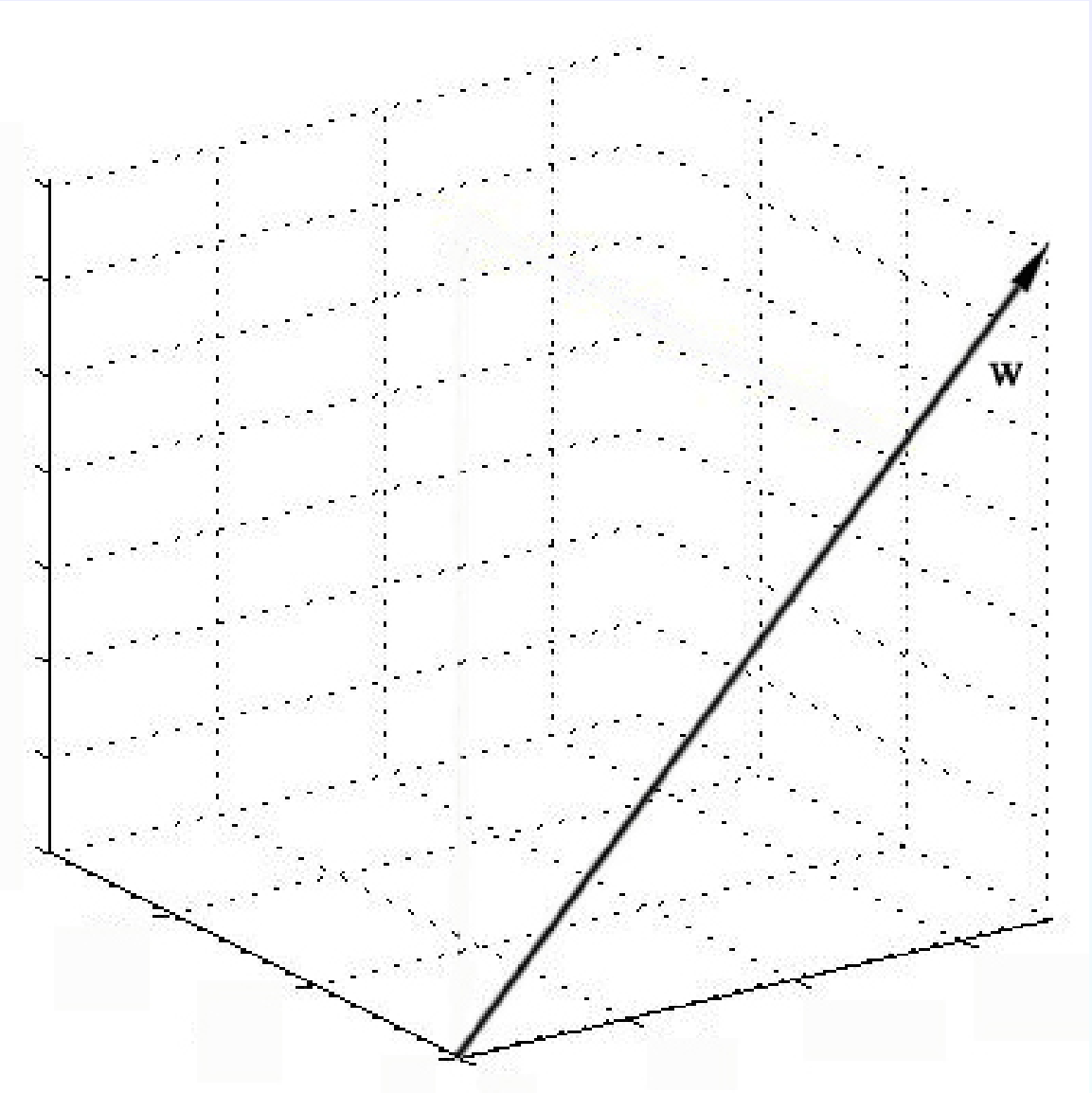
$$\frac{n(n-K)}{(n-1)K} \bar{y}^T S_y^{-1} \bar{y} > F_{(K,n-K),\alpha},$$

where $\bar{y}, S_y$ the sample mean and var-covar matrix of Y

- "search" throughout the entire multivariate space
- ineffective or inapplicable for **large K**, **small n** [3].

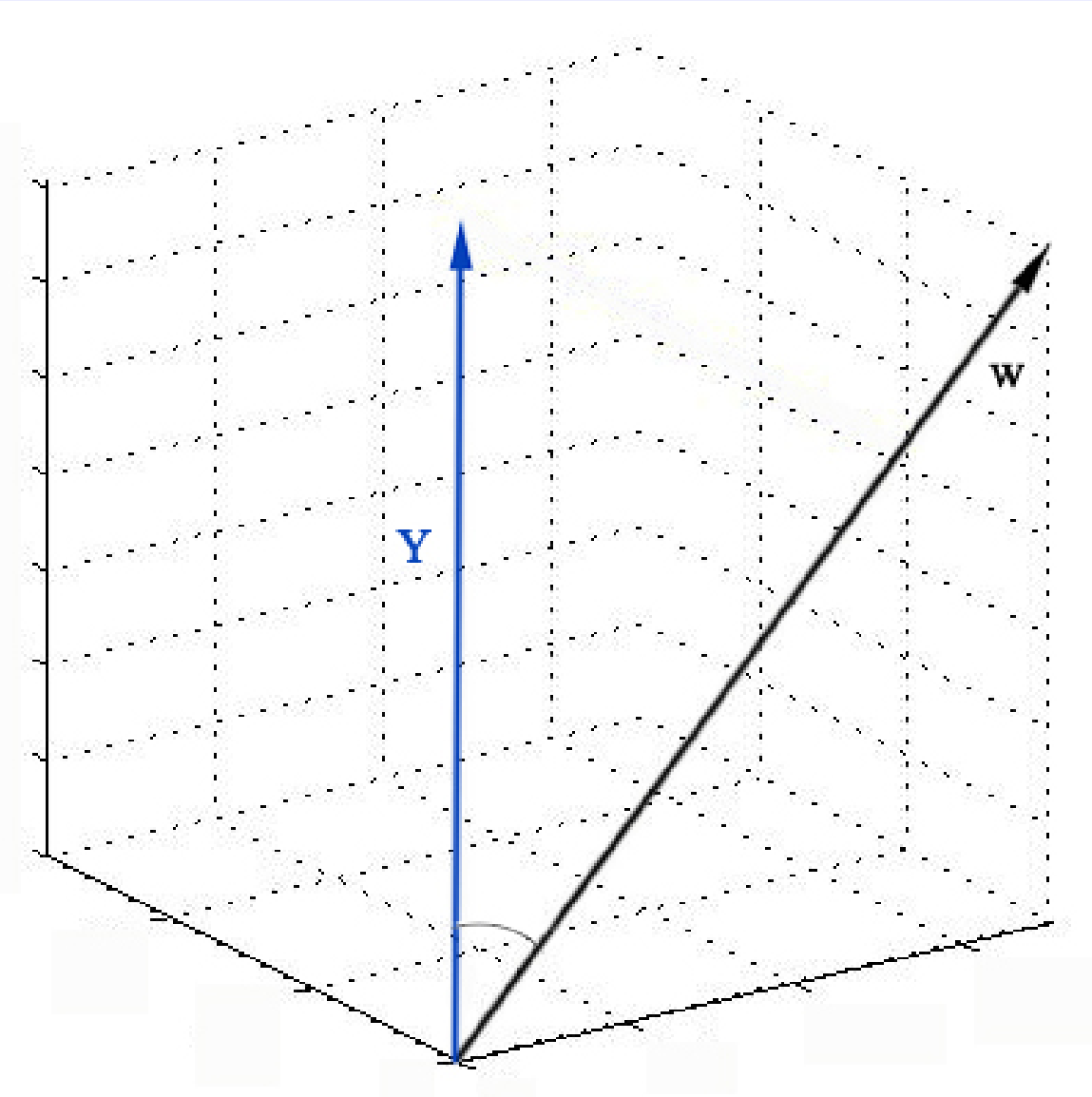
LINEAR COMBINATION (LC) TESTS

As opposed to searching throughout the entire multivariate space (T^2 test), LC tests search for effects only through a selected direction.



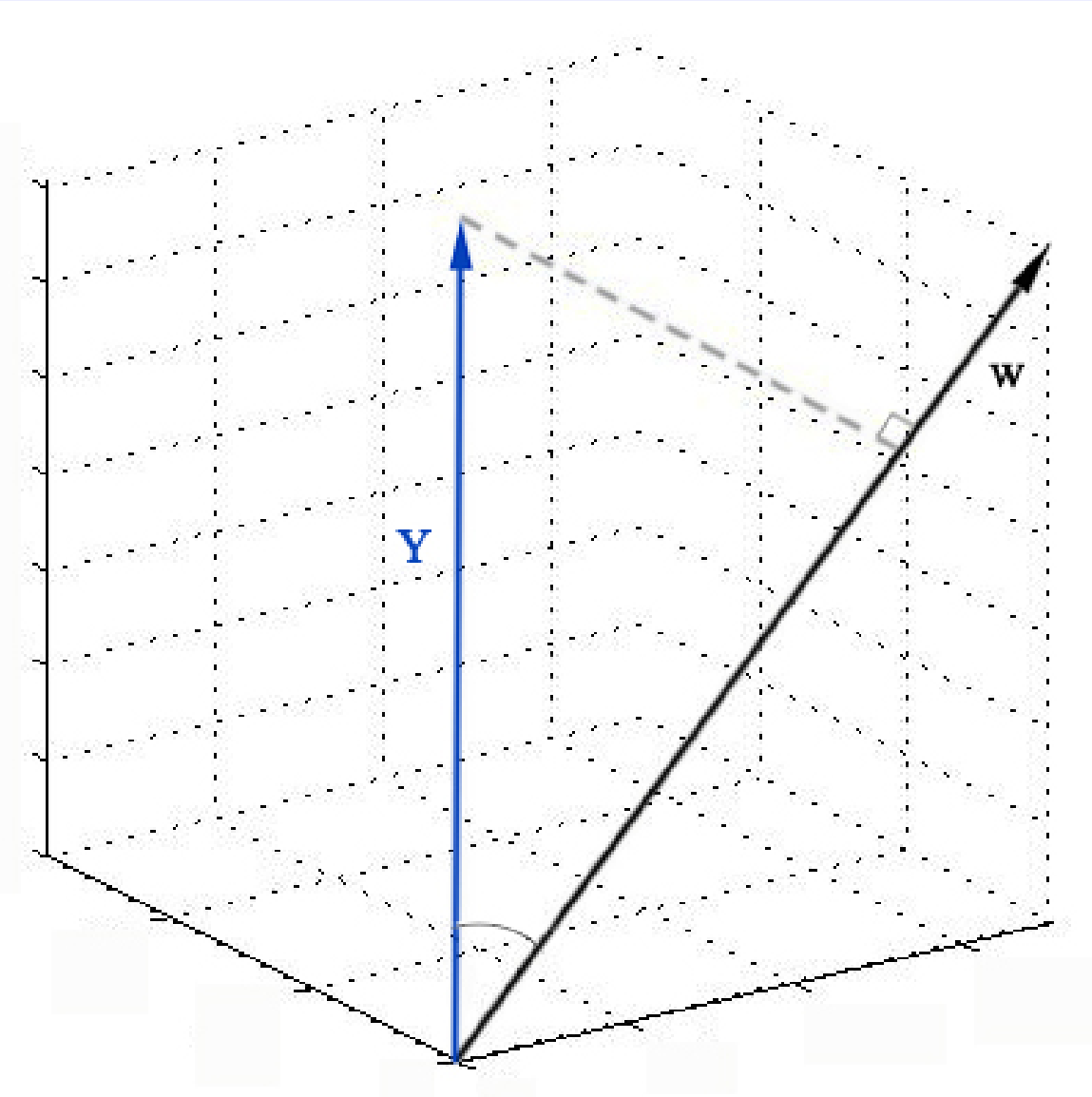
LINEAR COMBINATION (LC) TESTS

As opposed to searching throughout the entire multivariate space (T^2 test), LC tests search for effects only through a selected direction.



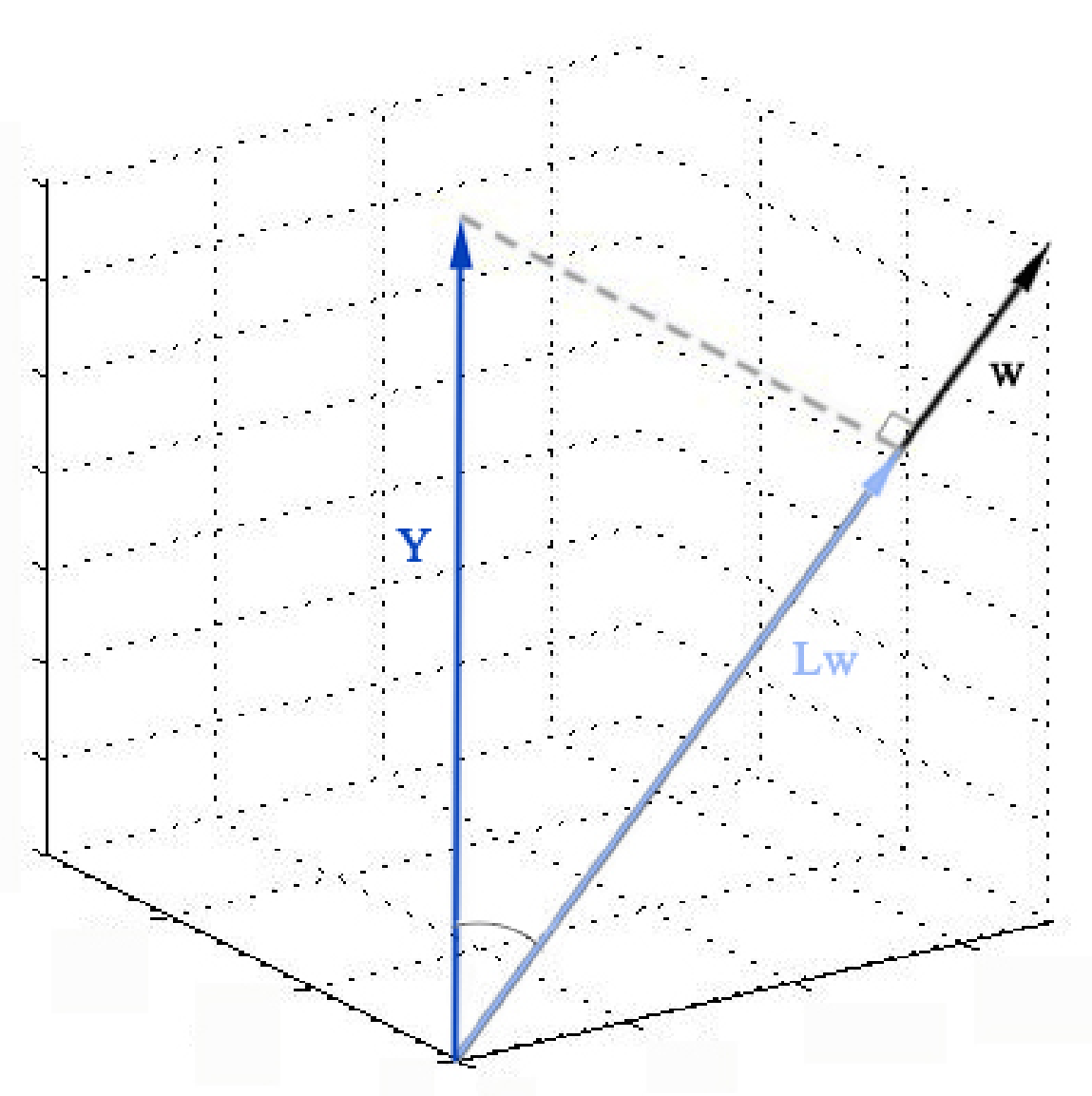
LINEAR COMBINATION (LC) TESTS

As opposed to searching throughout the entire multivariate space (T^2 test), LC tests search for effects only through a selected direction.



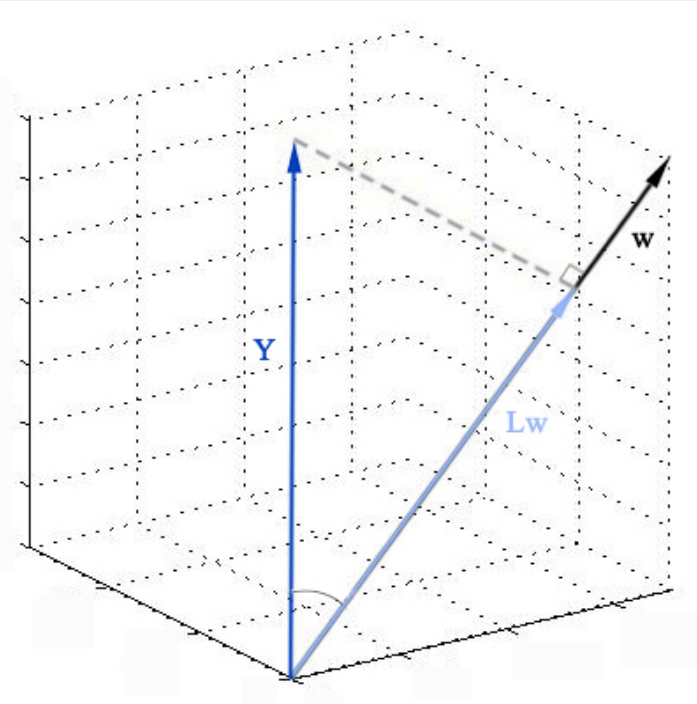
LINEAR COMBINATION (LC) TESTS

As opposed to searching throughout the entire multivariate space (T^2 test), LC tests search for effects only through a selected direction.



LINEAR COMBINATION (LC) TESTS

As opposed to searching throughout the entire multivariate space (T^2 test), LC tests search for effects only through a selected direction w.



In practice, we use the ***linear combination*** (projection magnitude)

$$L = w^T Y$$

to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

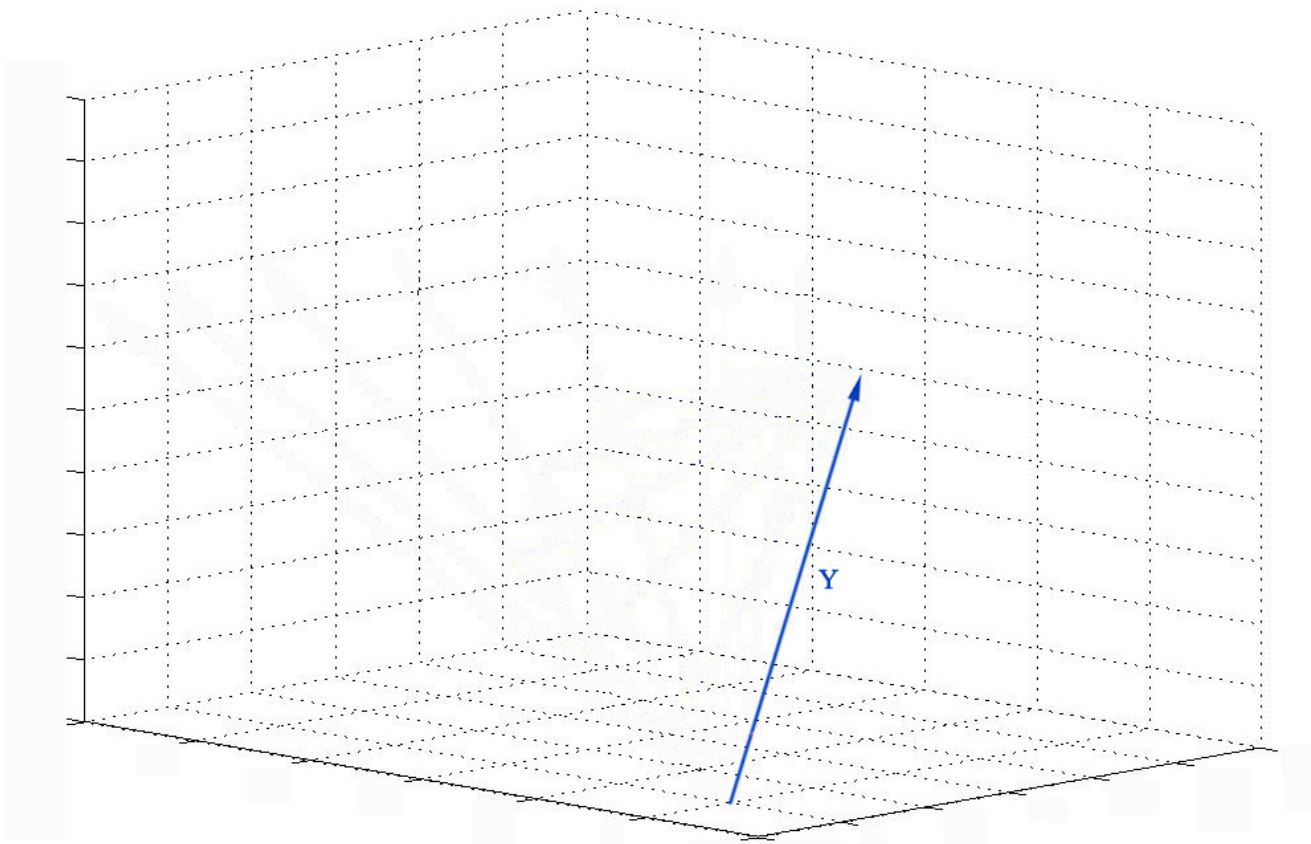
$\bar{L}$, σ^2, s^2: sample mean, variance and sample variance of L.

LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

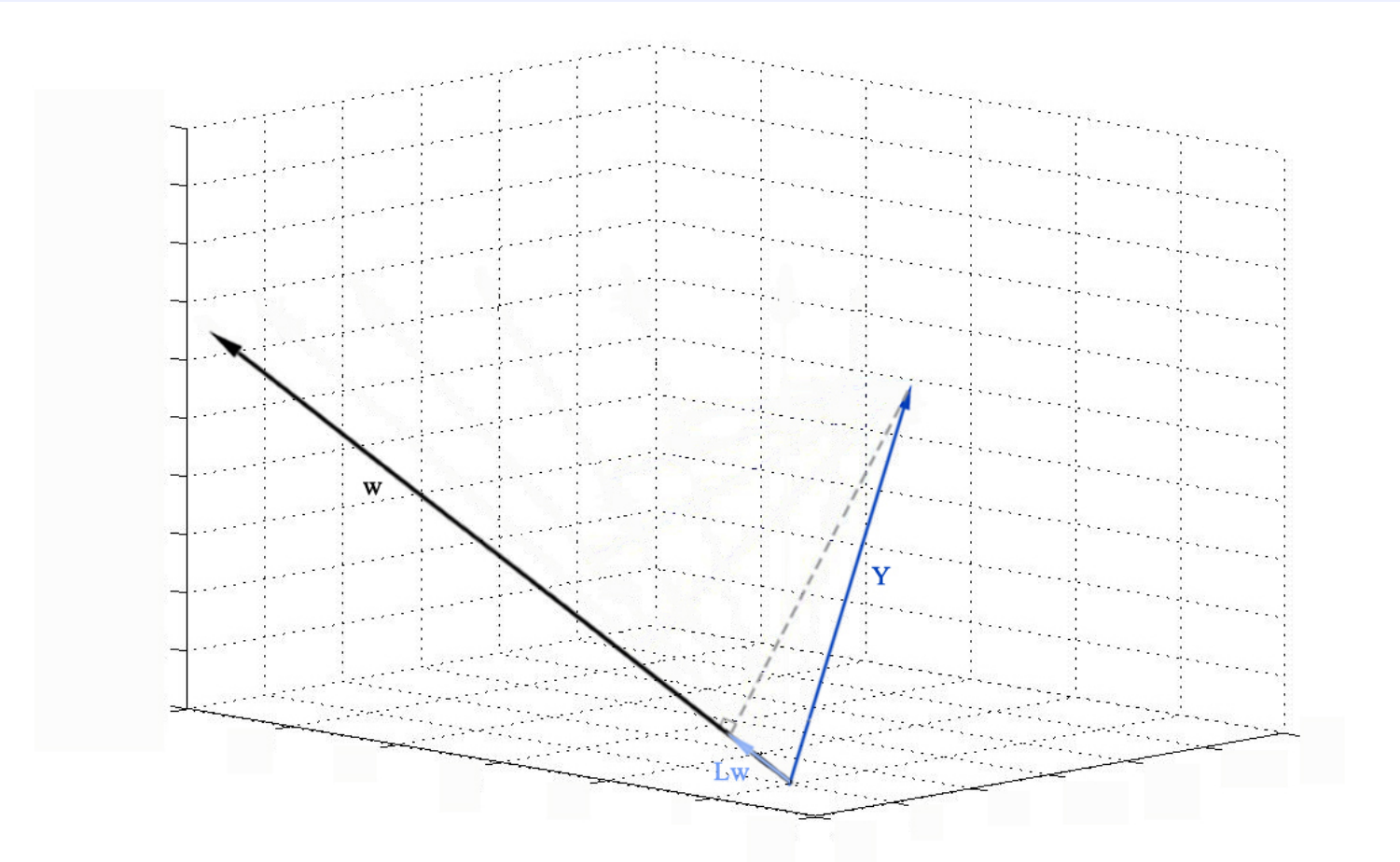


LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

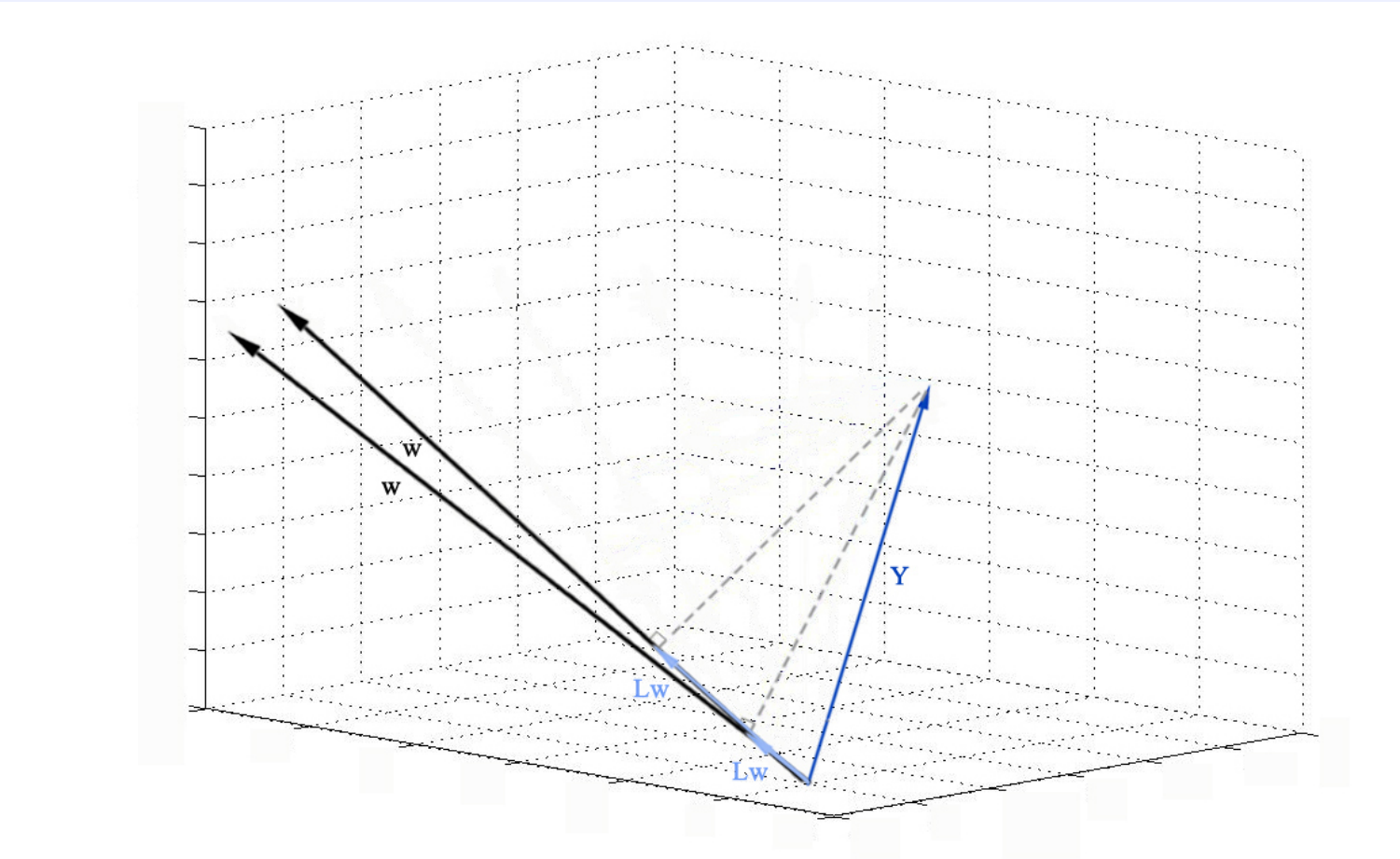


LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

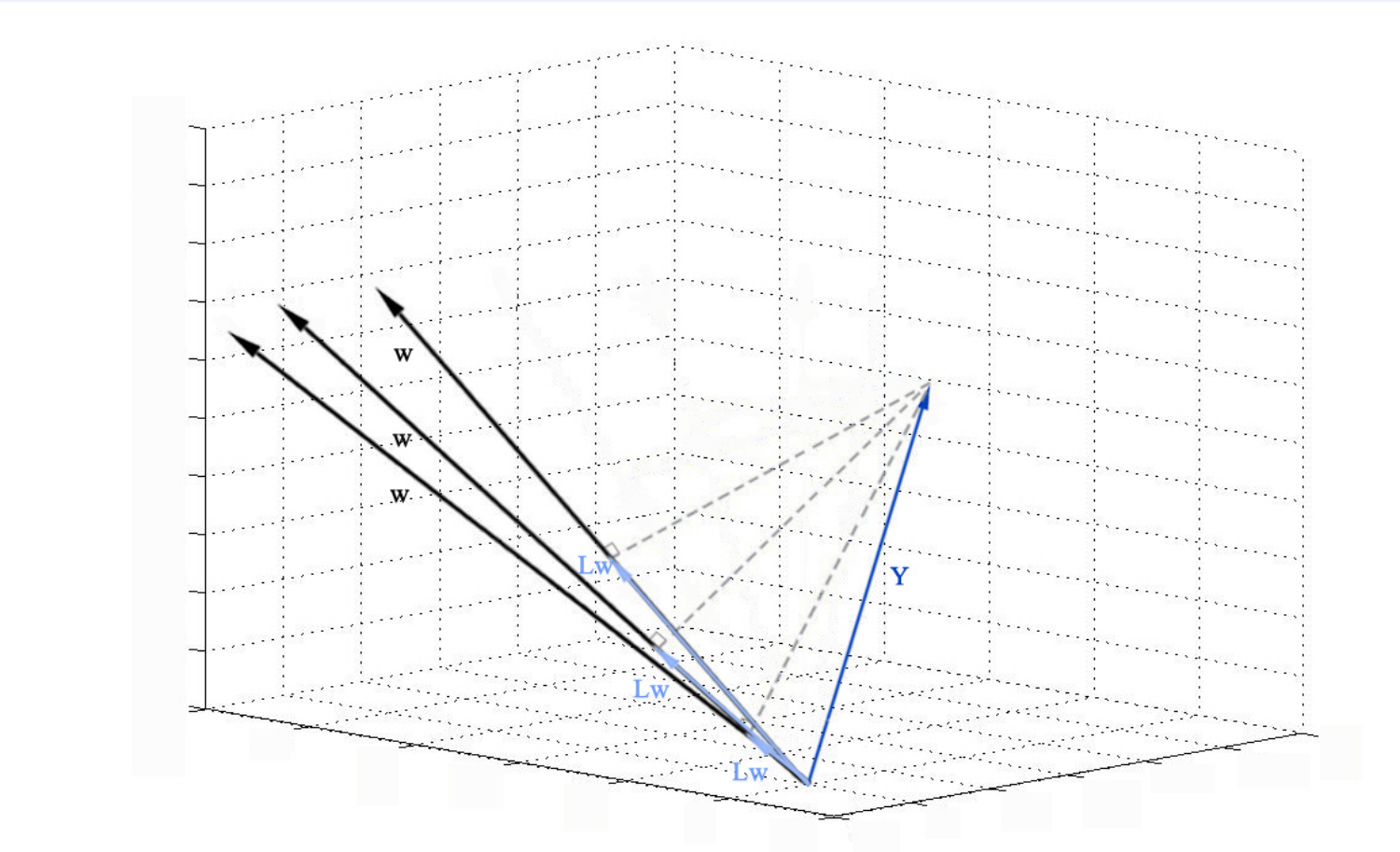


LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

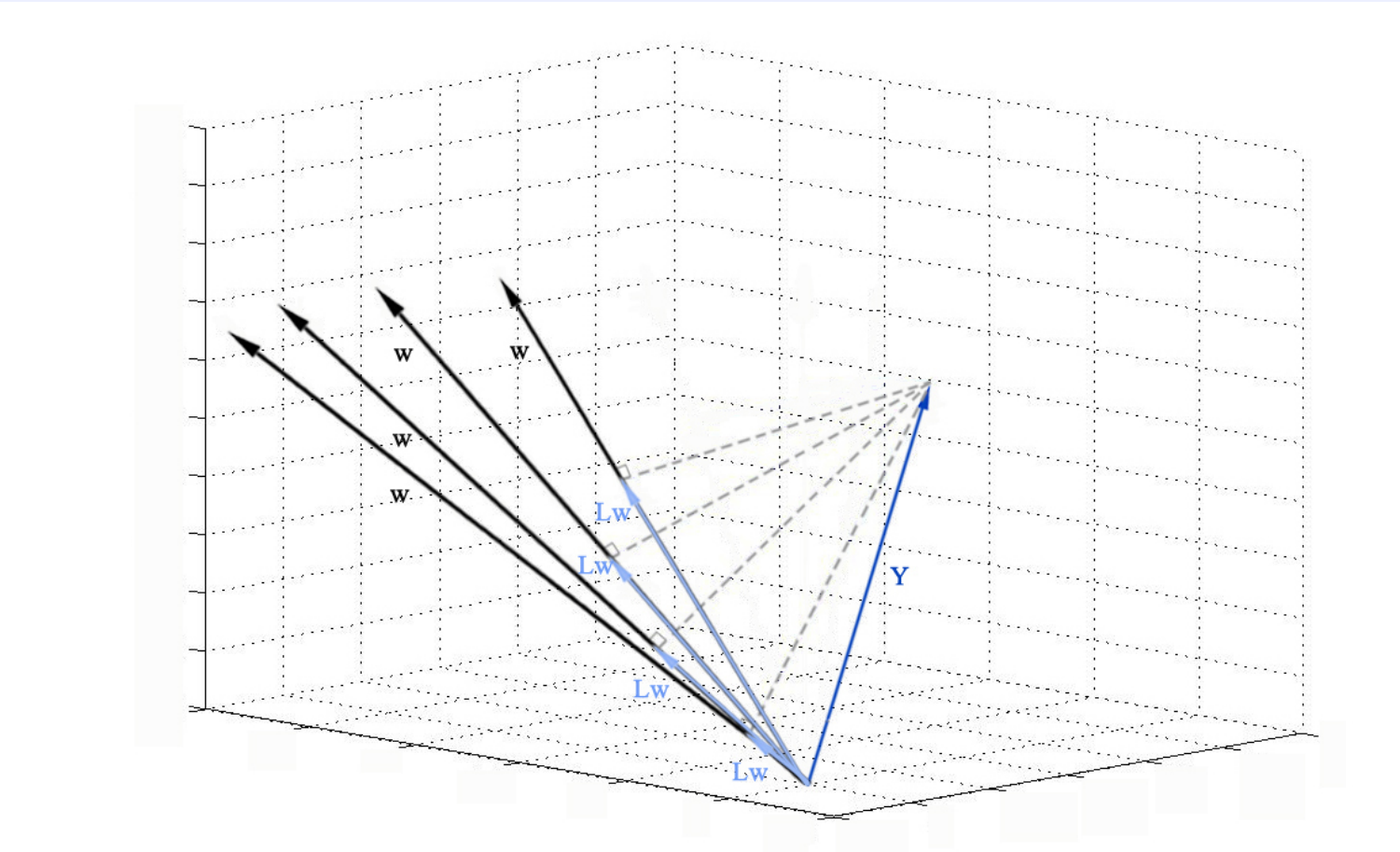


LINEAR COMBINATION (LC) TESTS

In practice, we use the linear combination $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

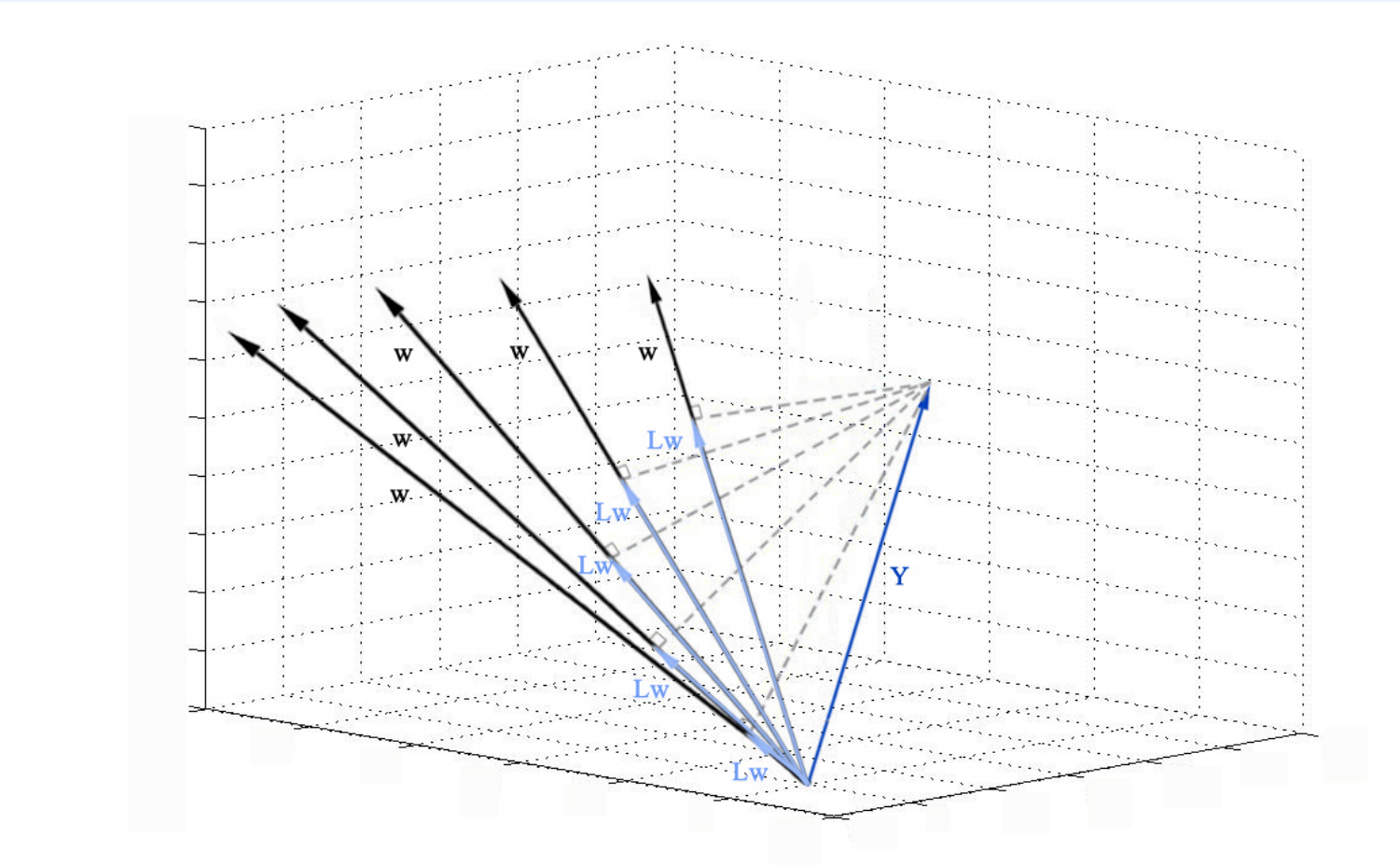


LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

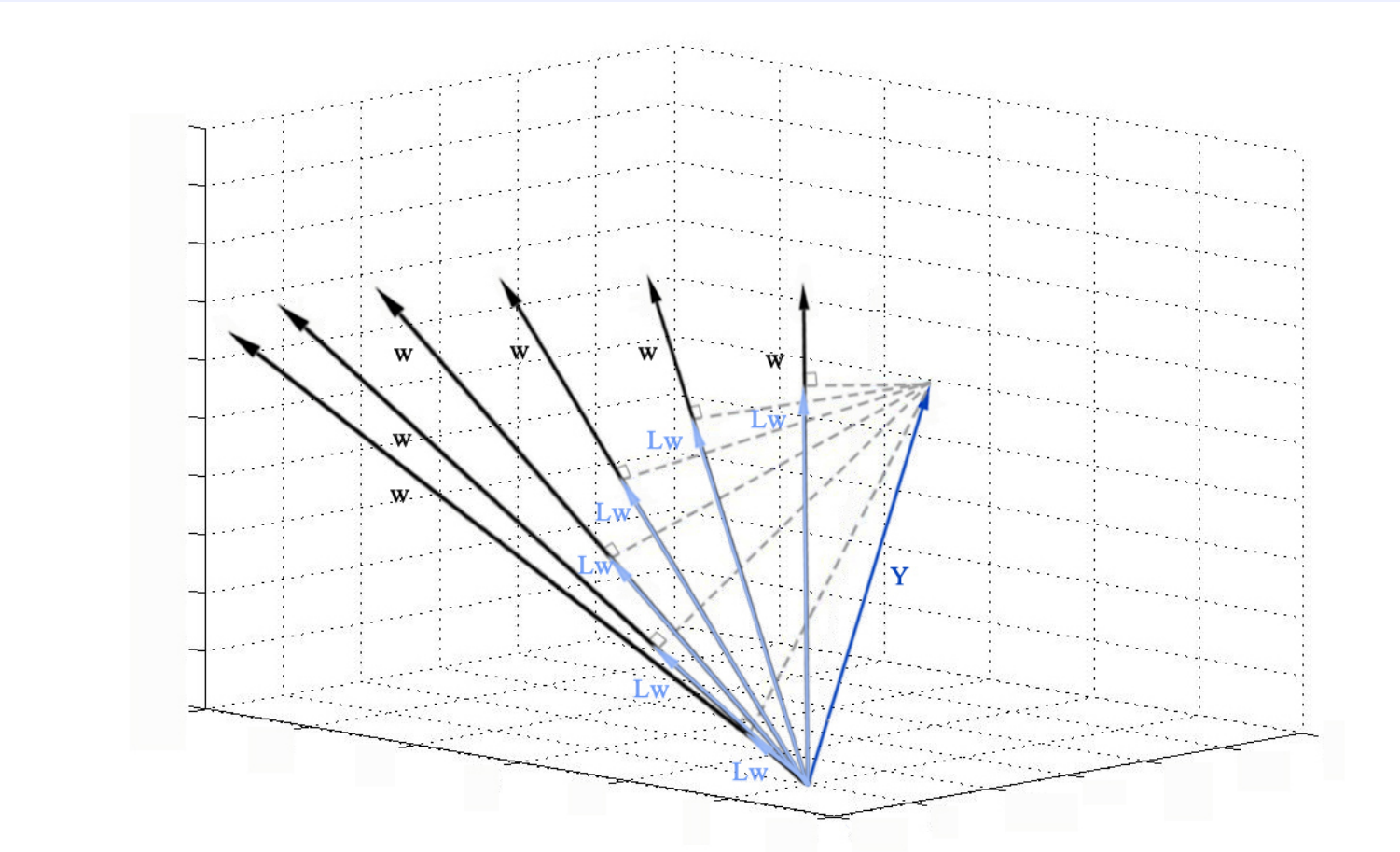


LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w

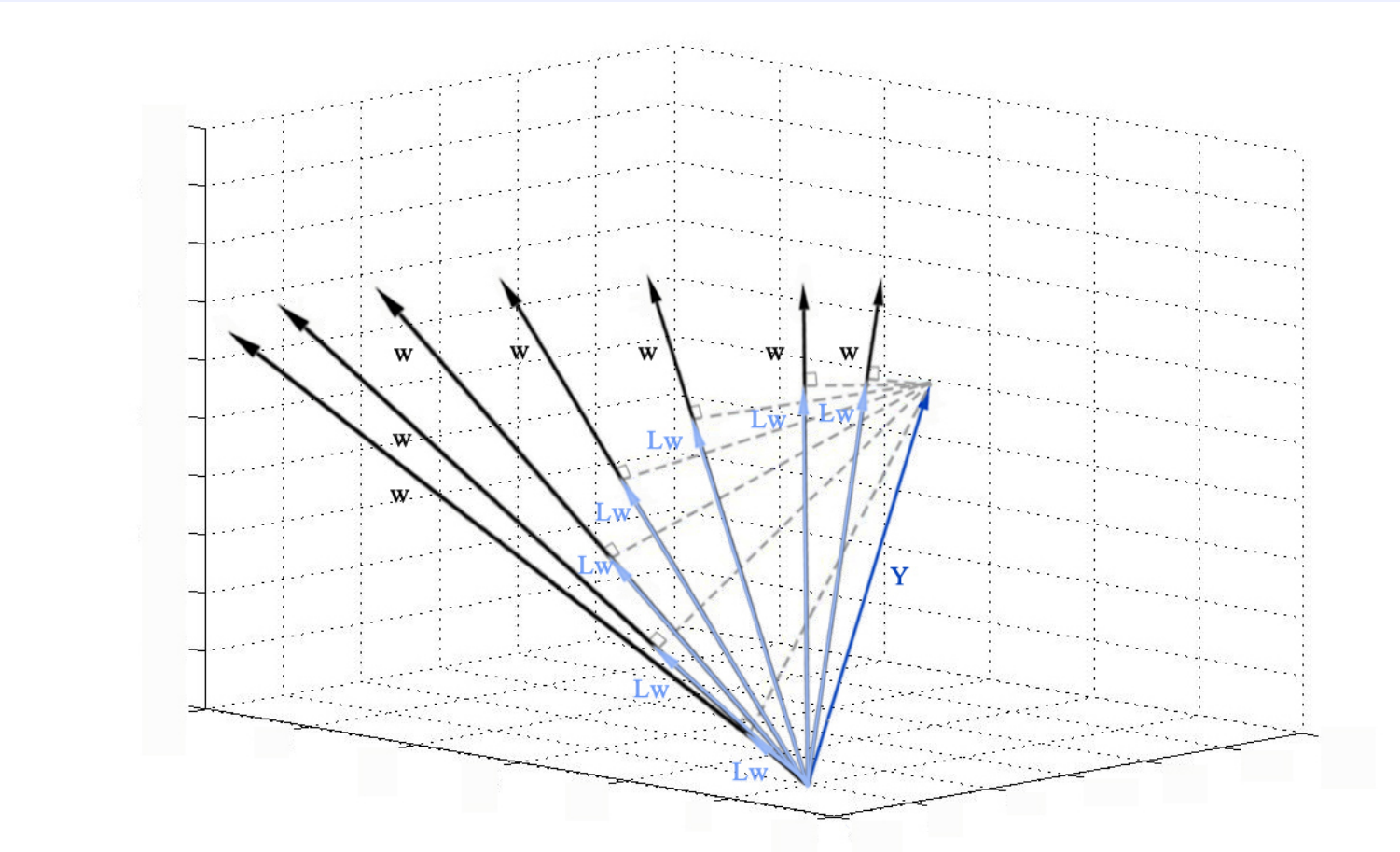


LINEAR COMBINATION (LC) TESTS

In practice, we use the *linear combination* $L = w^T Y$ to construct the z and t statistics

$$\Sigma \text{ known: } Z = \frac{\bar{L}}{\sigma/\sqrt{n}}, \qquad \Sigma \text{ unknown: } T = \frac{\bar{L}}{s/\sqrt{n}}$$

Crucial issue: The search direction w



THE SEARCH DIRECTION w

To see the importance of w, consider the distributions of the test statistics

$$Z \sim N(\delta\sqrt{n}, 1), \qquad T \sim t_{n-1}(\delta\sqrt{n}). \tag{1}$$

Here,

$$\delta = \frac{w^T\mu}{(w^T\Sigma w)^{1/2}} = \frac{\tilde{w}^T\tilde{\omega}_\star}{||\tilde{w}||} = ||\,\tilde{\omega}_\star\,|| \cos\left(ang(\tilde{w}, \tilde{\omega}_\star)\right), \tag{2}$$

where

$$\tilde{w} = \Sigma^{1/2}w: \quad \text{the transformed search direction}$$

and

$$\omega_\star = \Sigma^{-1}\mu, \quad \tilde{\omega}_\star = \Sigma^{1/2}\omega_\star. == \Sigma^{-1/2}\mu. \tag{3}$$

THE SEARCH DIRECTION w

To see the importance of w, consider the distributions of the test statistics

$$Z \sim N(\delta\sqrt{n}, 1), \qquad T \sim t_{n-1}(\delta\sqrt{n}). \tag{4}$$

Here,

$$\delta = \frac{w^T\mu}{(w^T\Sigma w)^{1/2}} = \frac{\tilde{w}^T\tilde{\omega}_\star}{\|\tilde{w}\|} = \|\tilde{\omega}_\star\| \cos\left(ang(\tilde{w}, \tilde{\omega}_\star)\right), \tag{5}$$

where

$$\tilde{w} = \Sigma^{1/2}w: \quad \text{the transformed search direction}$$

and

$$\omega_\star = \Sigma^{-1}\mu, \quad \tilde{\omega}_\star = \Sigma^{1/2}\omega_\star = \Sigma^{-1/2}\mu.. \tag{6}$$

THE SEARCH DIRECTION w

To see the importance of w, consider the distributions of the test statistics

$$Z \sim N(\delta\sqrt{n}, 1), \qquad T \sim t_{n-1}(\delta\sqrt{n}), \tag{7}$$

Here,

$$\delta = \frac{w^T\mu}{(w^T\Sigma w)^{1/2}} = \frac{\tilde{w}^T\tilde{\omega}_\star}{\|\tilde{w}\|} = \| \tilde{\omega}_\star \| \cos\left(ang(\tilde{w}, \tilde{\omega}_\star)\right), \tag{8}$$

where

$$\tilde{w} = \Sigma^{1/2}w: \quad \text{the transformed search direction}$$

and

$$\omega_\star = \Sigma^{-1}\mu, \quad \tilde{\omega}_\star = \Sigma^{1/2}\omega_\star = \Sigma^{-1/2}\mu. \tag{9}$$

Note that:

1. $\omega_\star$ defines the optimal search direction and ...
   - $\cos\left(ang(\tilde{w}, \tilde{\omega}_\star)\right)$ the **distance** of the search direction to the optimal
2. $\| \tilde{\omega}_\star \| = \left(\mu\Sigma^{-1}\mu\right)^{1/2}$: the Mahalanobis distance ...
   - that measures the **strength of the treatment effect**

PROPOSAL

- The optimal search direction $\omega_\star = \Sigma^{-1}\mu$ is *unknown*.

  ▶ **Question**: How to select w? Initially, using **prior information** and then by sequentially **adapting to accumulated data**

  - **Conduct an adaptive J-stage (here $J = 2$) study**
  - **Use predictive power to optimally derive w**

Other approaches include O'Brien [4] OLS LC z and t tests and Läuter [5] SS and PC LC t test [4].

FORMULATION: TWO-STAGE LC z AND t TESTS

- *Responses at j-th Stage*:

$$Y_{ji} \overset{\text{iid}}{\sim} N_K(\mu, \Sigma), \; i = 1, 2, ..., n_j, \; j = 1, 2 \tag{10}$$

- *Hypotheses*:

$$H_0\colon \mu = \underline{0} \text{ (no effect)} \quad \text{versus} \quad H_1\colon \mu \neq \underline{0} \tag{11}$$

- *Linear combination at j-th Stage*:

$$L_{ji} = w_j^T Y_{ji}, \; i = 1, 2, ..., n_j \; (w_j \neq \underline{0}) \tag{12}$$

- *Stage-wise z and t statistics*:

$$\Sigma \text{ known} : \; Z_j = \frac{\bar{L}_j}{\sigma_j/\sqrt{n_j}}, \qquad \Sigma \text{ unknown} : \; T_j = \frac{\bar{L}_j}{s_j/\sqrt{n_j}} \tag{13}$$

- *Stage-wise p-values*:

$$z\text{-test} : p_{z;j} = 2\Phi(-|Z_j|), \qquad t\text{-test} : p_{t;j} = 2\Psi_j(-|T_j|) \tag{14}$$

where $\Phi(\cdot)$, $\Psi_j(\cdot)$ the cdf's of $N(0,1)$ and t_{n_j-1}, respectively.

FORMULATION: TWO-STAGE LC z AND t TESTS 2

- ***Test***:

$$\textit{reject } \mathbf{H_0} \;\Leftrightarrow\; \{\mathbf{p_1} < \mathbf{a_1}\} \cup \{\mathbf{p_1} \in [\mathbf{a_1}, \mathbf{a_0}],\ \mathbf{p_1 p_2} < \mathbf{a_2}\} \tag{15}$$

Note: Fisher's product [6] $p_1 p_2$ is used for the final test (others are available)

- ***Power*** $(pr(\textit{reject } H_0))$:

$$\beta = pr(p_1 < a_1) + \int_{a_1}^{a_0} pr\left(p_1 p_2 < a_2 \mid p_1\right) g(p_1) dp_1, \tag{16}$$

where $g(\cdot)$ the density of p_1.

- ***Type I error rate*** $(pr(\textit{reject } H_0 \mid H_0 \textit{true}))$:

$$\alpha = pr_0(p_1 < a_1) + \int_{a_1}^{a_0} pr_0\left(p_1 p_2 < a_2 \mid p_1\right) g_0(p_1) dp_1 \tag{17}$$

Note: Equation (17) holds for w_2 depending on $\mathbf{y}_1$.

OPTIMAL SEARCH DIRECTION

It takes a few fairly easy derivations, to prove that the optimal, for the single stage LC tests, search direction $\omega_\star$ is also optimal for the two-stage LC test.

THEOREM 1

Under (10)*, the* ***power*** *of the two-stage z and t tests in* (16) *is maximized with respect to the weighting vectors* w_j, $j = 1, 2$, *if and only if the latter are both proportional to* $\omega_\star$ *in* (9).

- The weighting vector $\omega_\star$ depends on μ and Σ.
- We optimally use the available information to select w_j, $j = 1, 2$.

▶ *Optimality?* In terms of **predictive power**.

SELECTING SEARCH DIRECTION

- Predictive power [7]: $B_I = pr(\textit{reject } H_0 \mid I)$, I : information set
- I_0:
$$(\mu \mid \Sigma, I_0) \sim N_K\left(m_0, \Sigma/n_0\right), \qquad (\Sigma \mid I_0) \sim IW_{K\times K}\left(\nu_0, S_0^{-1}\right) \tag{18}$$
- $I_1 = \{I_0, y_1\}$

THEOREM 2

Under (10), (21), *the predictive power* $B_{I_{j-1}}$ *of the two-stage LC* $\mathbf{z}$ *test is maximized iff*
$$w_j \propto w_{z,j} = \Sigma^{-1} m_{j-1}, \quad j = 1, 2. \tag{19}$$
For $\nu_1 = \nu_0 + n_1 \to \infty$, *the predictive power* $B_{I_{j-1}}$ *of the two-stage LC* $\mathbf{t}$ *test is maximized iff*
$$w_j \propto w_{t,j} = S_{j-1}^{-1} m_{j-1}, \quad j = 1, 2. \tag{20}$$

$$m_1 = \frac{n_0 m_0 + n_1 \bar{y}_1}{n_0 + n_1}, \quad S_1 = S_0 + (n_1 - 1)S_{y_1} + \frac{n_0 n_1}{n_0 + n_1}(\bar{y}_1 - m_0)(\bar{y}_1 - m_0)^T$$

SELECTING SEARCH DIRECTION

- Predictive power [7]: $B_I = pr(\textit{reject } H_0 \mid I)$, I : information set
- I_0:
$$(\mu \mid \Sigma, I_0) \sim N_K\left(m_0, \Sigma/n_0\right), \qquad (\Sigma \mid I_0) \sim IW_{K\times K}\left(\nu_0, S_0^{-1}\right) \tag{21}$$
- $I_1 = \{I_0, y_1\}$

THEOREM 2

Under (10), (21), *the predictive power* $B_{I_{j-1}}$ *of the two-stage LC* **z** *test is maximized iff*
$$w_j \propto w_{z,j} = \Sigma^{-1} m_{j-1}, \quad j = 1, 2. \tag{22}$$
For $\nu_1 = \nu_0 + n_1 \to \infty$, *the predictive power,* B_{I_j} *of the two-stage LC* **t** *test is maximized iff*
$$w_j \propto w_{t,j} = S_{j-1}^{-1} m_{j-1}, \quad j = 1, 2. \tag{23}$$

Proposal: Use $w_{z,j}$ and $w_{t,j}$, $j = 1, 2$, to perform the two-stage LC z and t tests, respectively ▶ We will call these tests **adaptive z and t tests**

POWER ANALYSIS

We want to be able to explain power performance of the adaptive z and t tests, for a wide range of

1. the *critical values* a_0, a_1, a_2,
2. the prior, first stage and second stage *sample sizes* n_0, n_1 and n_2,
3. the *parameters* μ and Σ
4. the *prior estimates* m_0 (and S_0)

- Variables in 3, 4 are high-dimensional and complicatedly related to power.
- Can we find ***lower dimensional***, ***easily interpretable*** variables that are ***sufficient to describe power***?

POWER ANALYSIS: ADAPTIVE z TEST

Remark: In the **single-stage** LC z test, the mean of the z statistic is $\delta\sqrt{n}$ where

$$\delta = \| \tilde{\omega}_\star \| \cos ang\left(\tilde{w}, \tilde{\omega}_\star\right).$$

Thus, (a) the Mahalanobis distance $\| \tilde{\omega}_\star \| = \left(\mu\Sigma^{-1}\mu\right)^{1/2}$ and (b) the angle $ang\left(\tilde{w}, \tilde{\omega}_\star\right)$ (along with α and n) are sufficient to describe the power of the single-stage LC tests.

THEOREM 3

The values of (1) a_0, a_1, a_2, (2) n_0, n_1 *and* n_2, (3) $(\mu\Sigma^{-1}\mu)^{1/2}$ *and* (4) $ang\left(\tilde{w}_{z,1}, \tilde{\omega}_\star\right)$ *are sufficient to compute the power of the adaptive* z *test.*

Sketch of proof: First write power in terms of $(\tilde{\omega}_\star, \tilde{w}_{z,1})$ instead of (μ, Σ, m_0) and then show that power is invariant to rotations of $\tilde{w}_{z,1}$ around $\tilde{\omega}_\star$.

POWER ANALYSIS: ADAPTIVE t TEST

The prior estimates m_0 and S_0 are first "combined" to compute $w_{t,1}$, **but** then "split" to compute m_1 and S_1 to give $w_{t,2}$ ▶ *more complex situation*

Note that

$$\tilde{w}_{t,1} = \Sigma^{1/2} S_0^{-1} m_0 = D_1 \tilde{w}_{z,1},$$

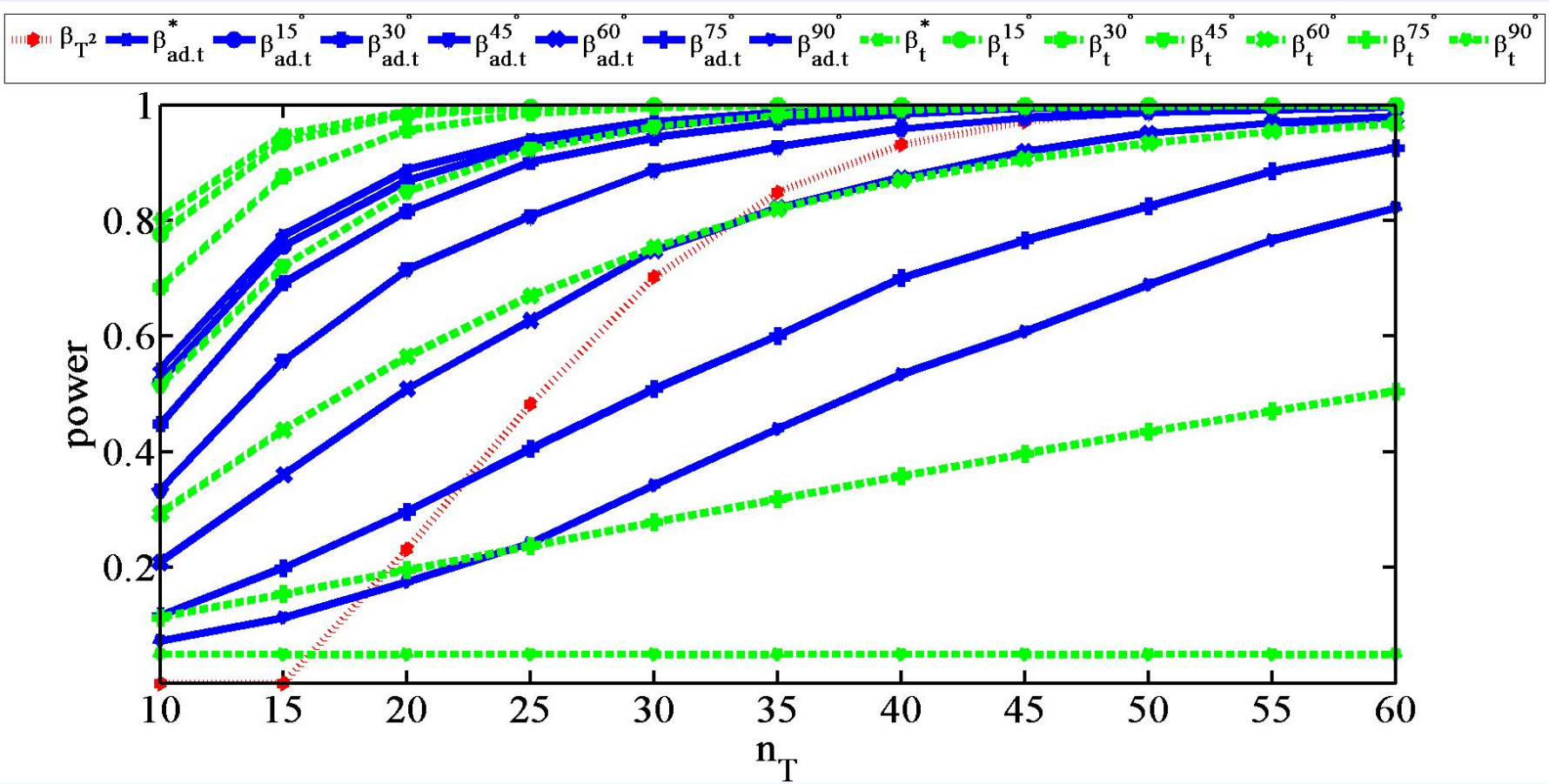
where $D_1 = \Sigma^{1/2} S_0^{-1} \Sigma^{1/2}$ the discrepancy between (Σ, S_0)
(and $\tilde{w}_{z,1}$: the selected w for known Σ)

THEOREM 4

The values of (1) α_0, α_1, a_2, (2) n_0, n_1 *and* n_2, (3) $(\mu\Sigma^{-1}\mu)^{1/2}$, (4) $(\lambda_k)_{k=1}^K$ (5) $(ang(\tilde{\omega}_\star, v_k))_{k=1}^K$, (6) $(ang(\tilde{w}_{z,1}, v_k))_{k=1}^K$, *where* $v_1, \ldots, v_K$ *the unit eigenvectors and* $\lambda_1, \ldots, \lambda_K$ *the corresponding eigenvalues of* D_1, *are sufficient to compute the power of the adaptive t test.*

Sketch of proof: First write power in terms of $(\tilde{\omega}_\star, \tilde{w}_{z,1}, D_1)$ instead of (μ, Σ, m_0, S_0) and then show that power is invariant to simultaneous rotations of the eigenvectors of D_1 and $\tilde{w}_{z,1}$ around $\tilde{\omega}_\star$.

POWER V TOTAL SAMPLE SIZE



- For small n_T, β_t^{ϕ} is larger (smaller) than $\beta_{ad.t}^{\phi}$ if ϕ is small (large)
- For large n_T, $\beta_{ad.t}$ reaches high levels even for $\tilde{w}_{z,1} \perp \tilde{\omega}_{\star}$ (unlike β_t)
- For increasing n_T, the angle ϕ that β_{T^2} surpasses $\beta_{ad.t}^{\phi}$, β_t^{ϕ} is decreasing

ROI ANALYSIS OF PHARMAFMRI DATA

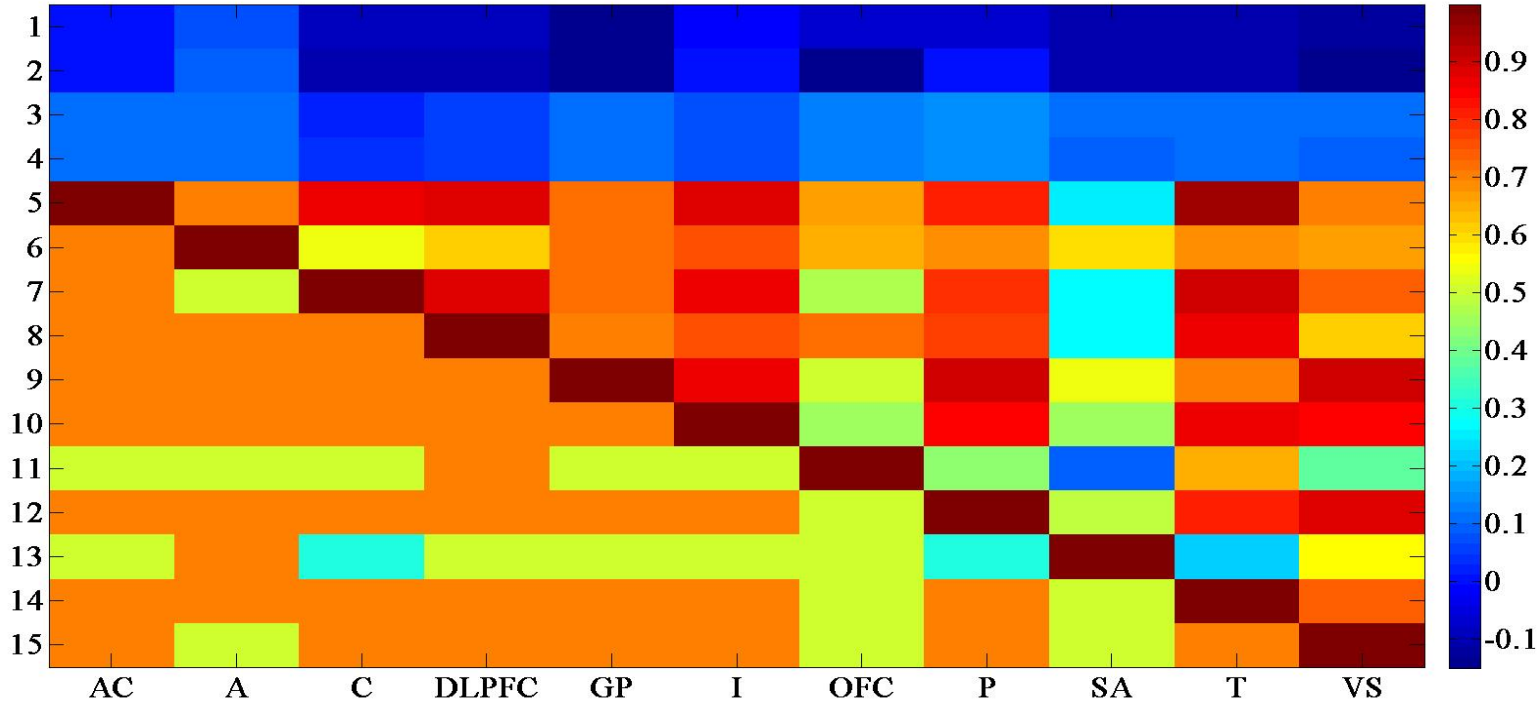


FIGURE: Means (l. 1), var's (l. 3) and corr's (upper triangle l. $5-15$) and their prior estimates (l. 2, 4 and lower triangle l. $5-15$) of ROI data of the sample ($n_T = 13$) of a GSK pharmafMRI study. The angle $ang(\tilde{\omega}_\star, \tilde{w}_{t,1}) = 67°$

$\beta_{ad.t} = 0.82$	$\beta_{T^2} = 0.30$	$\beta_{OLS,t} = 0.13$	$\beta_{SS,t} = 0.13$	$\beta_{PC,t} = 0.14$

TABLE: Power of T^2, adaptive, OLS (O'Brien), SS and PC (Läuter) LC t tests under the above estimates ($n_T = 13$, $\alpha = 0.05$)

SUMMARY

- We described a **two-step procedure** for ***formulating*** and ***testing*** the fundamental **global** null hypothesis of no treatment effect **in any of selected ROI**.
- **Step 1**: reduce fMRI data to ROI summary measures of treatment effect.
- **Step 2**: test the above global null hypothesis.
- We focused on Step 2 and propose **two-stage adaptive linear combination (LC) z and t tests**.
- We discussed the importance of the **search direction** w in these tests and we proposed an **optimal** method to derive w.
- We showed how to perform **power analysis** based on **low dimensional**, **easily interpretable** variables.
- We illustrated the advantages of our methods with respect to T^2 and alternative LC t tests for fairly precise prior information and for **large-K-small-n situation**.

REFERENCES

[1] G. D. Mitsis, G. D. Iannetti, T. S. Smart, I. Tracey, and R. G. Wise. Regions of interest analysis in pharmacological fMRI: How do the definition criteria influence the inferred result? Neuroimage, 40:121-132, 2007

[2] A. Dmitrienko A, A. C. Tamhane, and F. Bretz. Multiple Testing Problems in Pharmaceutical Statistics, Chapman & Hall/Crc Biostatistics Series, 2009

[3] T. W. Anderson. An introduction to multivariate statistical analysis, 2nd edition. John Wiley and Sons, 2003

[4] P. C. O'Brien. Procedures for comparing samples with multiple endpoints. Biometrics, 1984, 40, 1079-1087

[5] J. Läuter. Exact t and F tests for analyzing studies with multiple endpoints Biometrics, 1996, 52, 964-970

[6] Bauer, P. and Kohne, K. Evaluation of Experiments with Adaptive Interim Analyses. Biometrics, 1994, 50, 1029-1041

[7] D. J. Spiegelhalter, K. R. Abrams, and J. P. Myles. Bayesian approaches to clinical trials and health-care evaluation. John Wiley and Sons, 2004